Concave programmeringsproblemen

leren omgaan met de Stelling van Karush-Kuhn-Tucker

Pierre v. Mouche

Juni 2007

Verbeterde versie 0.8359 (maart 2008)

Inhoudsopgave

1 Inleiding **4**
1.1 Maximalisatie- en minimalisatieterminologie . . . 4
1.2 Maximalisatieproblemen . . . 4
1.3 Concave en niet-lineaire programmeringsproblemen . . . 5
1.4 Resultaten vooraf . . . 6

2 Zadelpunten **7**
2.1 Zadelpunten en lagrange-functie . . . 8
2.2 Zadelpunten versus maximaliseerders . . . 8

3 Differentieerbaarheid **11**
3.1 Karakterisatie van zadelpunten . . . 11
3.2 De Stelling van Karush-Kuhn-Tucker . . . 13

4 Opgaven **14**

A Appendix **15**
A.1 Existentie en uniciteit van maximaliseerders . . . 15
A.2 Convex gereedschap . . . 16

Voorwoord

Dit typoscript gaat over concave programmeringsproblemen (met expliciete ongelijkheids-, gelijkheids- en niet-negativiteitsrestricties). Er is ook wat aandacht voor de algemenere niet-lineaire programmeringsproblemen. Een hoofddoel is dat de lezer na het doorgewerkt te hebben goed weet hoe met behulp van de Stelling van Karush-Kuhn-Tucker concave programmeringsproblemen in geval van differentieerbare doelfunctie en restrictiefuncties aan te pakken. Om dat doel te bereiken is het niet per se nodig om de bewijzen te bestuderen. In de appendix wordt een en ander qua voorkennis in herinnering geroepen. Ook zijn er onder "convex gereedschap" wat resultaten uit de convexe analyse waarvan bij de bewijzen gebruik zal worden gemaakt. Ook als de lezer niet het fijne van deze resultaten weet, is het, omdat deze louter op een enkele plek toegepast worden, mogelijk de bewijzen in het typoscript door te werken. De appendix bevat tevens nog wat andere nuttige dingen. Diverse opgaven zijn opgenomen; voor wie daar nog niet genoeg aan heeft verwijst de auteur naar Bracken and McCormick (1968).

In paragraaf 1 wordt specifieke terminologie uitgelegd en de *setting* gefixeerd. De theorie die we daarvoor gaan opzetten maakt gebruik van de notie van zadelpunt. Paragraaf 2 is daaraan gewijd. Paragraaf 3 behelst de theorie voor het concave programmeringsprobleem met louter ongelijkheidsrestricties in geval van differentieerbare doel- en restrictiefuncties.

Na dit typoscript doorgewerkt te hebben, kan het interessant zijn het typoscript Balder (2003) te bestuderen voor een geavanceerdere presentatie van het concave programmeringsprobleem. Het geavanceerdere zit hem daar vooral in het loslaten van de differentieerbaarheidscondities.

Bij het vervaardigen van dit typoscript heeft de auteur vooral gebruikt gemaakt van het uitstekende boek Truchon (1987).

Uiteraard houdt de auteur zich aanbevolen voor op- en aanmerkingen die kunnen leiden tot verbeteringen.

<u>Notaties en conventies</u>

– In notaties als $\mathbb{R}^n$ en $\{1, \ldots, n\}$ is, tenzij anders vermeld, $n \geq 1$.

– Als we over "functie" spreken, dan is deze automatisch reëelwaardig en als we over "lineaire ruimte" spreken dan heeft deze automatisch $\mathbb{R}$ als scalairenlichaam.

– We gebruiken voor functies de volgende monotoniciteitstermen: dalend, strikt dalend, stijgend en strikt stijgend. De functie $f(x) := 7$ is zowel dalend als stijgend.

– Als we over "positief getal" spreken, dan bedoelen we daarmee een reëel getal groter dan 0.

– Een notatie als $f \restriction S$ duidt de beperking van de functie f tot de deelverzameling S van het domein van f aan.

– Elementen van $\mathbb{R}^n$ en afbeeldingen met waarden in $\mathbb{R}^n$ noteren we met vette symbolen; dus we schrijven bijvoorbeeld $\mathbf{x}$ voor een element van $\mathbb{R}^n$. En als $\mathbf{g}$ zo'n afbeelding is, dan noteren we met g_j haar projectie op de j-de coördinaat.

– Voor matrixproducten gebruiken we het symbool $\star$.

– Elementen van $\mathbb{R}^n$ vatten we op als rijvectoren. Dit heeft als belangrijk voordeel dat rij-vector-notaties zoals $\mathbf{x} = (x_1, \ldots, x_n)$ typografisch doorgaans prettiger zijn dan kolom-vector-notaties, maar een nadeel is dat we dan getransponeerden $\mathbf{x}^t$ moeten gebruiken in uitdrukkingen zoals $A \star \mathbf{x}^t$ waar A een $m \times n$-matrix is.

1 Inleiding

1.1 Maximalisatie- en minimalisatieterminologie

Zij X een verzameling en $f : X \to \mathbb{R}$ een functie. Een element a van X heet

- maximaliseerder van f als $f(x) \leq f(a)$ $(x \in X)$. We zeggen dan nog dat f een maximum in a heeft en dat $f(a)$ het maximum van f is. We noteren dat maximum met $\max(f)$. De verzameling der maximaliseerders noteren we met $\operatorname{argmax}(f)$.
- minimaliseerder van f als $f(x) \geq f(a)$ $(x \in X)$. We zeggen dan nog dat f een minimum in a heeft en dat $f(a)$ het minimum van f is. We noteren dat minimum met $\min(f)$. De verzameling der minimaliseerders noteren we met $\operatorname{argmin}(f)$.
- extremaliseerder van f als a een minimaliseerder of maximaliseerder van f is. We zeggen dan nog dat f een extremum in a en dat $f(a)$ een extremum van f is.

Soms laat men aan al deze noties nog het bijvoeglijk naamwoord "globaal" vooraf gaan, zo dat men dus spreekt van bijvoorbeeld globaal minimum en van globale maximaliseerder. Dit doet men om het onderscheid te benadrukken met de lokale varianten van deze noties welke als volgt gedefinieerd zijn.

Zij X een topologische ruimte, $f : X \to \mathbb{R}$ een functie. Een element a van X heet

- lokale maximaliseerder van f als er een omgeving U van a bestaat zodanig dat a een maximaliseerder van de functie $f \upharpoonright U$ is. We zeggen dan nog dat f een lokaal maximum in a heeft en dat $f(a)$ een lokaal maximum van f is.
- lokale minimaliseerder van f als er een omgeving U van a bestaat zodanig dat a een minimaliseerder van de functie $f \upharpoonright U$ is. We zeggen dan nog dat f een lokaal minimum in a heeft en dat $f(a)$ een lokaal minimum van f is.
- lokaal extreem van f als a een lokaal maximum of lokaal minimum van f is. We zeggen dan nog dat f een lokaal extreem in a heeft en dat $f(a)$ een lokaal extreem van f is.

Opmerkingen:

- $\operatorname{argmax}(f) = \operatorname{argmin}(-f)$ en als $\max(f)$ wel-gedefinieerd is, dan is $\min(-f)$ dat ook en geldt $\max(f) = -\min(-f)$.
- Er zijn generalisaties van bovenstaande noties in verschillende richtingen mogelijk.
  - Voor $f : X \to \mathbb{R}^p$. Het is de theorie van de vectormaximalisatie die zich daarmee bezig houdt.
  - Voor $f : X \to \bar{\mathbb{R}}$ waar $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.
- Als $X \subseteq \mathbb{R}^n$, $a \in X$ en $f : X \to \mathbb{R}$ is, dan hebben we hierboven onder meer gedefinieerd wat het betekent dat a een lokale maximaliseerder van f is door gebruik te maken van de topologie op X. Er geldt in dat verband ook nog het volgende: a is een lokale maximaliseerder van f d.e.s.d.a. er is een omgeving U van a in $\mathbb{R}^n$ zodanig dat a een maximaliseerder van $f \upharpoonright X \cap U$ is.

1.2 Maximalisatieproblemen

Gegeven een functie $f : X \to \mathbb{R}$, waar X een verzameling is, kunnen we het volgende probleem bekijken: bepaal, als dat bestaat, het maximum $\max(f)$ van f, en in dat geval ook de verzameling $\operatorname{argmax}(f)$ der maximaliseerders.[1] We noteren dit probleem met

$$\mathrm{MAX}\, f,$$

[1] $\sup(f)$, i.e. het supremum van f is altijd wel-gedefinieerd, maar $\max(f)$ hoeft dat niet te zijn.

en spreken van het maximalisatieprobleem voor f. f heet ook wel de doelfunctie. Als $\max(f)$ bestaat, dan verstaan we onder $\max(f)$ tezamen met $\operatorname{argmax}(f)$ de oplossing van $\mathrm{MAX}\, f$. Het broertje van dit probleem betreft het minimalisatieprobleem

$$\mathrm{MIN}\, f$$

voor f het om het minimum $\min(f)$ en de verzameling der minimaliseerders en $\operatorname{argmin}(f)$ gaat.

Opgemerkt zij dat als we het maximalisatieprobleem van f begrijpen we ook het minimalisatieprobleem van f begrijpen en omgekeerd (door op $-f$ over te stappen). Wij maken hier de keuze dat we ons daarom veelal tot maximalisatieproblemen zullen bespreken.

In de praktijk gaat het werk bij een maximalisatieprobleem doorgaans zitten in het bepalen van de maximaliseerders en volgt het maximum dan simpelweg door f in één der maximaliseerders uit te rekenen.

Het domein van de functie f hierboven is dus X. We noteren dat ook wel expliciet als volgt in het maximalisatieprobleem:

$$\mathrm{MAX}_{x \in X}\ f(x).$$

Bekijken we nu het maximalisatieprobleem van een functie $f : X \to \mathbb{R}$ onder de restrictie dat alleen elementen van X die in een of andere deelverzameling B van X liggen in aanmerking komen. We noteren zo'n probleem ook wel als

$$\mathrm{MAX}_{\text{o.d.r. } x \in B}\ f(x),$$

of uitgebreider als

$$\mathrm{MAX}_{\substack{x \in X \\ \text{o.d.r. } x \in B}}\ f(x).$$

Hier staat "o.d.r." voor "onder de restrictie". We noemen in zo'n probleem de elementen uit B nog de toelaatbare elementen. Het gaat daarbij natuurlijk om niets anders dan het maximalisatieprobleem $\mathrm{MAX}\ f \restriction B$. Maar, zoals we zullen zien, blijkt het mogelijk te zijn voor speciale klassen van dergelijke maximalisatieproblemen de extra structuur van zo'n B uit te buiten en zo tot meer op maat gesneden theorieën te komen. B behelst expliciete restricties, als U zo wilt.

1.3 Concave en niet-lineaire programmeringsproblemen

Er bestaan diverse op maat gesneden theorieën voor het maximalisatieprobleem

$$\mathrm{MAX}_{\substack{x \in X \\ \text{o.d.r. } x \in B}}\ f(x).$$

Deze verlaten bovenstaande algemene *setting*. Een goed ontwikkelde theorie bestaat voor het geval dat X een deelverzameling van de $\mathbb{R}^n$ is en de deelverzameling B van X gegeven wordt als de oplossingsverzameling van een eindig aantal ongelijkheids- en gelijkheidsrestricties op X.[2] De volgende definitie verschaft de meest belangrijke gevallen. In deze definitie gebruiken we voor de hand liggende vectornotaties.

Definitie 1 Een niet-lineair programmeringsprobleem is een maximalisatieprobleem van de vorm

$$\mathrm{MAX}_{\substack{\mathbf{x} \in X \\ \text{o.d.r. } \mathbf{g}^K(\mathbf{x}) \leq \mathbf{d}^K,\ \mathbf{g}^L(\mathbf{x}) = \mathbf{d}^L,\ \mathbf{x}^I \geq \mathbf{0}}}\ f(\mathbf{x})$$

waar X een deel van $\mathbb{R}^n$ is, waar K, L disjuncte delen zijn van $M := \{1, \ldots, m\}$ met $K \cup L = M$, I een deel van $N := \{1, \ldots, n\}$ is, $f : X \to \mathbb{R}$ is, $g_i : X \to \mathbb{R}$ $(i \in M)$ en $d_i \in \mathbb{R}$ $(i \in M)$. In het geval dat X convex, f concaaf, g_i $(i \in K)$ convex en g_i $(i \in L)$ affien zijn, spreken we van een concaaf programmeringsprobleem.[3] En als f, g_i $(i \in M)$ affien zijn van een lineair programmeringsprobleem. $\diamond$

[2] X wordt daarbij vaak open verondersteld.

[3] En hier voor convexe f en MAX vervangende door MIN van een convex programmeringsprobleem. Merk in dat verband nog op dat concave functies "gemaximaliseerd willen worden" en convexe functies "geminimaliseerd willen worden" (als U begrijpt wat de auteur bedoelt). Daarom zijn convexe programmeringsproblemen natuurlijker in het geval van minimalisatie.

Let goed op: de g_j zijn convex bij een concaaf programmeringsprobleem.

We zullen vaak het symbool $\spadesuit$ gebruiken om een niet-lineair-programmeringsprobleem met notaties zoals in Definitie 1 mee aan te duiden.

Het onderwerp van dit typoscript zijn de concave programmeringsproblemen en in mindere mate de niet-lineaire; de lineaire komen helemaal niet expliciet aan de orde.

Vaak bekijkt men in de literatuur situaties waar $L = \emptyset$, i.e. met louter ongelijkheidsrestricties. Dat doet men omdat gelijkheidsrestricties theoretische complicaties veroorzaken. Natuurlijk kan men een gelijkheidsrestrictie behapstukken met twee ongelijkheidsrestricties: een gelijkheidsrestrictie van de vorm $g_i(\mathbf{x}) = d_i$ equivalent is met de twee ongelijkheidsrestricties $g_i(\mathbf{x}) \leq d_i$ en $-g_i(\mathbf{x}) \leq -d_i$. Op die manier krijgt men dan weliswaar wel een niet-lineair programmeringsprobleem met louter ongelijkheidsrestricties, maar diverse resultaten voor niet-lineaire programmerings met louter gelijkheidsrestricties zijn op dat probleem niet meer van toepassing (omdat de resterende ongelijkheidsrestricties "verdorven" zijn). Vandaar dat we bijzondere aandacht zullen hebben voor niet-lineaire-programmeringsproblemen met louter ongelijkheidsrestricties, i.e. in geval $L = \emptyset$ en $I = N$; z.b.d.a. nemen we daar ook $d_i = 0$ $(i \in M)$:[4]

$$\text{MAX} \underset{\text{o.d.r. } g_i(\mathbf{x}) \leq 0 \ (1 \leq i \leq m), \ x_i \geq 0 \ (1 \leq i \leq n)}{\mathbf{x} \in X} f(\mathbf{x}).$$

We zullen naar zo'n probleem verwijzen met het symbool $\spadesuit_{\leq}$.

Voor niet-lineaire programmeringsproblemen met louter gelijkheidsresctricties raadplege men de literatuur over de Multiplicatorenregel van Lagrange.

Er zijn verhandelingen waar men ervoor kiest om steeds $X = \mathbb{R}^n$ te nemen. Dit vinden wij minder geslaagd omdat er genoeg functies zijn die niet (zonder meer) op de hele $\mathbb{R}^n$ gedefinieerd zijn. Verder merk op dat we voorlopig niets over X veronderstellen, in het bijzondere (nog) niet dat X open is;[5] dat gebeurt pas als we differentieerbaarheidseigenschappen van de f en g_i erbij gaan betrekken.

De verzameling der toelaatbare elementen van een niet-lineair programmeringsprobleem kan best leeg zijn. Maar dat geval is niet bijster interessant.

1.4 Resultaten vooraf

Zonder nog noemenswaardige theorie ontwikkeld te hebben, geven we in deze deelparagraaf alvast wat nuttige resultaten.

Propositie 1 *Beschouw een concaaf programmeringsprobleem. Zij B de verzamelingen der toelaatbare elementen.*

1. *B is convex.*
2. *De verzameling der maximaliseerders is convex.*
3. *Indien $f \upharpoonright B$ strikt quasi-concaaf is, bestaat de verzameling der maximaliseerders uit hoogstens één punt.*
4. *Indien B compact en niet-leeg is en $f \upharpoonright B$ continu is, dan is de verzameling der maximaliseerders niet-leeg en compact.* $\diamond$

Bewijs.— 1. Stel $\mathbf{x}, \mathbf{y}$ zijn toelaatbare elementen en laat $\lambda, \mu \in [0,1]$ met $\lambda + \mu = 1$. Dan $\mathbf{g}^K(\mathbf{x}) \leq \mathbf{d}^K$, $\mathbf{g}^L(\mathbf{x}) = \mathbf{d}^L$, $\mathbf{x}^I \geq \mathbf{0}$ en $\mathbf{g}^K(\mathbf{y}) \leq \mathbf{d}^K$, $\mathbf{g}^L(\mathbf{y}) = \mathbf{d}^L$, $\mathbf{y}^I \geq \mathbf{0}$. Daaruit volgt voor $i \in K$ omdat g_i convex is

$$g_i(\lambda\mathbf{x} + \mu\mathbf{y}) \leq \lambda g_i(\mathbf{x}) + \mu g_i(\mathbf{y}) \leq \lambda d_i + \mu d_i = d_i$$

[4] Restricties van de vorm $g_i(\mathbf{x}) \leq d_i$ (en $g_i(\mathbf{x}) = d_i$) kan men behapstukken door dergelijke restricties met $d_i = 0$ door over te stappen op $g_i - d_i$.

[5] Maar natuurlijk in het geval van een concaaf programmeringsprobleem is X tevens convex.

en dus $\mathbf{g}^K(\lambda\mathbf{x}+\mu\mathbf{y}) \leq \mathbf{d}^K$. En voor $i \in L$ omdat g_i affien is

$$g_i(\lambda\mathbf{x}+\mu\mathbf{y}) = \lambda g_i(\mathbf{x}) + \mu g_i(\mathbf{y}) = \lambda d_i + \mu d_i = d_i$$

en dus $\mathbf{g}^L(\lambda\mathbf{x}+\mu\mathbf{y}) = \mathbf{d}^L$. En voor alle $i \in I$ geldt $(\lambda\mathbf{x}+\mu\mathbf{y})_i \geq \mathbf{0}$.

2. Dat volgt uit 1 en ☺ 2 (uit de appendix).
3. Vanwege ☺ 1.
4. Vanwege de Stelling van Weierstrass, i.e. ☺ 4. Q.E.D.

Stelling 1 *Beschouw het niet-lineaire programmeringsprobleem*

$$\underset{\substack{\mathbf{x}\in X \\ o.d.r.\ \mathbf{g}^K(\mathbf{x})\leq\mathbf{d}^K,\ \mathbf{g}^L(\mathbf{x})=\mathbf{d}^L,\ \mathbf{x}^I\geq\mathbf{0}}}{\mathrm{MAX}} f(\mathbf{x})$$

waar X een convex deel van $\mathbb{R}^n$ is en de $g_i : X \to \mathbb{R}$ $(i \in M)$ quasi-convex zijn.

Stel $\mathbf{a}$ is een inwendig punt van X dat toelaatbaar is, met zelfs $\mathbf{g}^K(\mathbf{a}) = \mathbf{d}^K$, en waarin f en elke g_m differentieerbaar is, f pseudo-concaaf is[6] *en*

$$(\nabla f)(\mathbf{a}) - \sum_{i\in M} \lambda_m \nabla g_i(\mathbf{a}) = \mathbf{0} \textit{ voor zekere } \lambda_i \geq 0.$$

Dan is $\mathbf{a}$ een maximaliseerder. ⋄

Bewijs.— Fixeer een toelaatbaar element $\mathbf{y} \in X$. We laten zien dat $f(\mathbf{y}) \leq f(\mathbf{a})$. Merk op dat $g_i(\mathbf{y}) \leq d_i$ en $g_i(\mathbf{a}) = d_i$. Omdat g_i quasi-convex is, volgt voor elke $r \in [0,1]$ dat $g_i(\mathbf{a}+r(\mathbf{y}-\mathbf{a})) \leq \max(g_i(\mathbf{a}), g_i(\mathbf{y})) = g_i(\mathbf{a})$. Met andere woorden: "als men zich vanuit $\mathbf{a}$ in de richting $\mathbf{y}-\mathbf{a}$ beweegt, dan neemt g_i niet toe". Dus is $\mathbf{y}-\mathbf{a}$ geen stijgrichting van g_i in $\mathbf{a}$ en moet[7]

$$\nabla g_i(\mathbf{a}) \cdot (\mathbf{y}-\mathbf{a}) \leq 0.$$

Omdat $\nabla f(\mathbf{a})\cdot(\mathbf{y}-\mathbf{a}) = (\sum_{i\in M} \lambda_i \nabla g_i(\mathbf{a}))\cdot(\mathbf{y}-\mathbf{a}) = \sum_{i\in M} \lambda_i (\nabla g_i(\mathbf{a})\cdot(\mathbf{y}-\mathbf{a}))$ en $\lambda_i \geq 0$ $(i \in M)$, volgt

$$\nabla f(\mathbf{a}) \cdot (\mathbf{y}-\mathbf{a}) \leq 0.$$

De pseudoconcaviteit van f te $\mathbf{a}$ impliceert nu tenslotte dat $f(\mathbf{y}) \leq f(\mathbf{a})$. Q.E.D.

Corollarium 1 *Beschouw het niet-lineaire programmeringsprobleem*

$$\underset{\substack{\mathbf{x}\in\mathbb{R}^n_+ \\ o.d.r.\ g_i(\mathbf{x})\leq 0\ (1\leq i\leq m)}}{\mathrm{MAX}} f(\mathbf{x}).$$

waar de $g_i : \mathbb{R}^n_+ \to \mathbb{R}$ $(1 \leq i \leq m)$ quasi-convex zijn. Stel $\mathbf{a} \in \mathbb{R}^n_{++}$ is toelaatbaar met zelfs $g_i(\mathbf{a}) = 0$ $(1 \leq i \leq m)$, f en elke g_i is te $\mathbf{a}$ differentieerbaar, f is te $\mathbf{a}$ pseudo-concaaf en

$$(\nabla f)(\mathbf{a}) - \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{a}) = \mathbf{0} \textit{ voor zekere } \lambda_i \geq 0.$$

Dan is $\mathbf{a}$ een maximaliseerder. ⋄

2 Zadelpunten

In deze paragraaf gaat het om het niet-lineaire programmeringsprobleem

$$\underset{\substack{\mathbf{x}\in X \\ \text{o.d.r. } \mathbf{g}^K(\mathbf{x})\leq\mathbf{d}^K,\ \mathbf{g}^L(\mathbf{x})=\mathbf{d}^L,\ \mathbf{x}^I\geq\mathbf{0}}}{\mathrm{MAX}} f(\mathbf{x}). \qquad \spadesuit$$

We introduceren voor dit maximalisatieprobleem de notie van zadelpunt en lagrange-functie en bestuderen het verband tussen zadelpunten en maximaliseerders.

[6] Dat laatste is volgens ☺ 9 zeker zo als f continu en quasi-concaaf is en de afgeleide van f in $\mathbf{a}$ ongelijk aan nul is.

[7] Ter herinnering: Zij $X \subseteq \mathbb{R}^n$, $f : X \to \mathbb{R}$ en $\mathbf{a}$ een inwendig punt van X. Als f in $\mathbf{a}$ differentieerbaar is, dan is elke vector $\mathbf{d} \in \mathbb{R}^n$ waarvoor $Df(\mathbf{a})\mathbf{d} < 0$ een afdaalrichting van f in a, i.e. dan bestaat er een $\delta > 0$ zodanig dat $f(\mathbf{a}+\lambda\mathbf{d}) < f(\mathbf{a})$ $(0 < \lambda < \delta)$.

2.1 Zadelpunten en lagrange-functie

Definitie 2 Beschouw het niet-lineaire programmeringsprobleem ♠. Onder de lagrange-functie (geassocieerd met ♠) verstaat men de functie $\mathcal{L} : X \times \mathbb{R}^m \to \mathbb{R}$ gedefinieerd door

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) := f(\mathbf{x}) + \sum_{i \in M} \lambda_i (d_i - g_i(\mathbf{x})).$$

De λ_i heten nog lagrange-multiplicatoren.[8] $\diamond$

Merk op dat elke functie $\mathcal{L}(\mathbf{x}, \cdot)$ affien is en dat in geval van een concaaf programmeringsprobleem voor elke $\boldsymbol{\lambda} \in \mathbb{R}^m$ met $\boldsymbol{\lambda}^K \geq \mathbf{0}$ de functie $\mathcal{L}(\cdot, \boldsymbol{\lambda})$ concaaf is. Merk ook op dat voor elke toelaatbare $\mathbf{x}$ en $\boldsymbol{\lambda} \in \mathbb{R}^m$ met $\boldsymbol{\lambda}^K \geq 0$

$$\sum_{i \in M} \lambda_i (d_i - g_i(\mathbf{x})) \geq 0.$$

Definitie 3 Beschouw het niet-lineaire programmeringsprobleem ♠ en zij $\mathcal{L} : X \times \mathbb{R}^m \to \mathbb{R}$ de ermee geassocieerde lagrange-functie. Zij

$$Y := \{(\mathbf{x}, \boldsymbol{\lambda}) \in X \times \mathbb{R}^m \mid \mathbf{x}^I \geq \mathbf{0} \text{ en } \boldsymbol{\lambda}^K \geq \mathbf{0}\}.$$

Een vector $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \in Y$ heet zadelpunt van $\mathcal{L}$ (in Y) als

$$\mathcal{L}(\mathbf{x}, \bar{\boldsymbol{\lambda}}) \leq \mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \leq \mathcal{L}(\bar{\mathbf{x}}, \boldsymbol{\lambda}) \quad ((\mathbf{x}, \boldsymbol{\lambda}) \in Y). \ \diamond$$

Merk op dat in geval van $♠_{\leq}$ er geldt dat

$$\mathcal{L} : X \times \mathbb{R}^m \to \mathbb{R},$$

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}),$$

$$Y = (X \cap \mathbb{R}^n_+) \times \mathbb{R}^m_+,$$

en dat in dat geval $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}})$ een zadelpunt van $\mathcal{L}$ is, dan en slechts dan als $\bar{\mathbf{x}}$ een maximaliseerder is van de functie $\mathcal{L}(\cdot, \bar{\boldsymbol{\lambda}})$ op $X \cap \mathbb{R}^n_+$ en $\bar{\boldsymbol{\lambda}}$ een minimaliseerder is van de functie $\mathcal{L}(\bar{\mathbf{x}}, \cdot)$ op $\mathbb{R}^m_+$.

2.2 Zadelpunten versus maximaliseerders

Propositie 2 *Beschouw het niet-lineaire programmeringsprobleem ♠. Als $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}})$ een zadelpunt van de lagrange-functie $\mathcal{L}$ is, dan is $\bar{\mathbf{x}}$ een maximaliseerder en geldt $\mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = f(\bar{\mathbf{x}})$.* $\diamond$

Bewijs.— We hebben voor alle $\boldsymbol{\lambda} \in \mathbb{R}^m$ met $\boldsymbol{\lambda}^K \geq \mathbf{0}$

$$\mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \leq \mathcal{L}(\bar{\mathbf{x}}, \boldsymbol{\lambda}).$$

Dit is equivalent met

$$\sum_{i \in M} \bar{\lambda}_i (d_i - g_i(\bar{\mathbf{x}})) \leq \sum_{i \in M} \lambda_i (d_i - g_i(\bar{\mathbf{x}})). \tag{1}$$

Door $\boldsymbol{\lambda} = \mathbf{0}$ te nemen, volgt

$$\sum_{i \in M} \bar{\lambda}_i (d_i - g_i(\bar{\mathbf{x}})) \leq 0.$$

We hebben dus

$$\mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = f(\bar{\mathbf{x}}),$$

[8] Synoniem: lagrange-vermenigvuldigers.

i.e.
$$\sum_{i \in M} \bar{\lambda}_i(d_i - g_i(\bar{\mathbf{x}})) = 0 \tag{2}$$
bewezen als we kunnen laten zien dat veronderstellen van de strikte ongelijkheid
$$\sum_{i \in M} \bar{\lambda}_i(d_i - g_i(\bar{\mathbf{x}})) < 0$$
een tegenspraak oplevert. Veronderstel dan nu dat deze geldt. Dan volgt dat $\bar{\lambda}_{i_0} > 0$ en $d_{i_0} - g_{i_0}(\bar{\mathbf{x}}) < 0$ voor tenminste één $i_0 \in M$ of $\bar{\lambda}_{i_0} < 0$ en $d_{i_0} - g_{i_0}\bar{\mathbf{x}}) > 0$ voor tenminste één $i_0 \in M$. Dit impliceert dat men $\sum_{i \in M} \lambda_i(d_i - g_i(\bar{\mathbf{x}})) < \sum_{i \in M} \bar{\lambda}_i(d_i - g_i(\bar{\mathbf{x}}))$ kan krijgen door in het eerste geval λ_{i_0} voldoende groot te nemen en voldoende klein in het tweede geval. Maar dat is een tegenspraak met (1).

We hebben nu volgens (1) dus $\sum_{i \in M} \lambda_i(d_i - g_i(\bar{\mathbf{x}})) \geq 0$ voor alle $\boldsymbol{\lambda} \in \mathbb{R}^m$ met $\boldsymbol{\lambda}^K \geq \mathbf{0}$. Deze ongelijkheid impliceert dat $\mathbf{g}^L(\bar{\mathbf{x}}) = \mathbf{d}^L$ en $\mathbf{g}^K(\bar{\mathbf{x}}) \leq \mathbf{d}^K$. Dus $\bar{\mathbf{x}}$ is toelaatbaar.

Vanwege (2) wordt de ongelijkheid $\mathcal{L}(\mathbf{x}, \bar{\boldsymbol{\lambda}}) \leq \mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}})$ $(\mathbf{x}^I \geq \mathbf{0})$:
$$f(\mathbf{x}) + \sum_{i \in M} \bar{\lambda}_i(d_i - g_i(\mathbf{x})) \leq f(\bar{\mathbf{x}}) \ (\mathbf{x}^I \geq \mathbf{0}).$$
Voor een toelaatbare $\mathbf{x}$ geldt $\mathbf{x}^I \geq \mathbf{0}$ en $\sum_{i \in M} \bar{\lambda}_i(d_i - g_i(\mathbf{x})) \geq 0$ en volgt daarom $f(\mathbf{x}) \leq f(\bar{\mathbf{x}})$. Q.E.D.

De volgende stelling laat zien dat voor een concaaf programmeringsprobleem als alles even meezit ook de omkering van Propositie 2 geldt.

Propositie 3 *Stel* $\bar{\mathbf{x}}$ *is een maximaliseerder van het concave programmeringsprobleem* ♠. *Stel ook dat aan de volgende* <u>*conditie van Slater*</u> *is voldaan: er bestaat een* $\mathbf{a} \in X$ *met*

- $\mathbf{g}^K(\mathbf{a}) \ll \mathbf{d}^K$, $\mathbf{g}^L(\mathbf{a}) = \mathbf{d}^L$ *en* $\mathbf{a}^I \gg \mathbf{0}$*;*
- *de familie van vectoren* $((\nabla g_i)(\mathbf{a}))_{i \in L}$ *is linear onafhankelijk.*

Dan is er een $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^m$ *met* $\bar{\boldsymbol{\lambda}}^K \geq \mathbf{0}$ *zodanig dat* $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}})$ *een zadelpunt van de lagrangefunctie* $\mathcal{L}$ *is en*
$$\mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = f(\bar{\mathbf{x}})$$
is. ◇

Bewijs.— Zet $X' := \{\mathbf{x} \in X \mid \mathbf{x}^I \geq \mathbf{0}\}$. X' is convex. Omdat $\bar{\mathbf{x}}$ een maximaliseerder is, geldt dat elke $\mathbf{x} \in X'$ die aan $\mathbf{d}^K - \mathbf{g}^K(\mathbf{x}) \geq \mathbf{0}$ en $\mathbf{d}^L - \mathbf{g}^L(\mathbf{x}) = \mathbf{0}$ voldoet ook aan $f(\mathbf{x}) - f(\bar{\mathbf{x}}) \leq 0$ voldoet. Toepassing van ☺ 7(1) levert het bestaan van $(\bar{\boldsymbol{\mu}}, \bar{v}) \in (\mathbb{R}^m \times \mathbb{R}) \setminus \{\mathbf{0}\}$ zodanig dat $(\bar{\boldsymbol{\mu}}^K, \bar{v}) \geq \mathbf{0}$ en
$$\bar{v}(f(\mathbf{x}) - f(\bar{\mathbf{x}})) + \sum_{i \in M} \bar{\mu}_i(d_i - g_i(\mathbf{x})) \leq 0 \ \ (\mathbf{x} \in X'). \tag{3}$$
En met ☺ 7(2) zien we dat
$$\sum_{i \in M} \bar{\mu}_i(d_i - g_i(\bar{\mathbf{x}})) = 0, \tag{4}$$
dus $\mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\mu}}) = f(\bar{\boldsymbol{x}})$.

We bewijzen nu uit het ongerijmde dat $\bar{v} > 0$. Stel dus $\bar{v} \leq 0$; nu zelfs
$$\bar{v} = 0.$$
Indien $\bar{\boldsymbol{\mu}}^K > \mathbf{0}$, dan zou volgen dat $\sum_{i \in M} \bar{\mu}_i(d_i - g_i(\mathbf{a})) > 0$ en daarmee een tegenspraak met (3). Dus geldt, omdat $\bar{\boldsymbol{\mu}} \geq \mathbf{0}$
$$\bar{\boldsymbol{\mu}}^K = \mathbf{0}.$$

In geval $K = M$ geeft dat een tegenspraak met $(\bar{\mu}, \bar{v}) \neq \mathbf{0}$. Maar voor het geval waar dat niet per se zo is, is een verdere redenering nodig. Welnu, omdat $(\bar{\boldsymbol{\mu}}, \bar{v}) \neq \mathbf{0}$ volgt $\bar{\boldsymbol{\mu}}_L \neq \mathbf{0}$. Er is daarom een $k \in L$ met $\bar{\mu}_k > 0$. omdat elke g^i $(i \in L)$ affien is, is er een $\#L \times n$-matrix A en $\mathbf{b} \in \mathbb{R}^L$ zodanig dat $g^L(\mathbf{x}) = \mathbf{b} + A \times \mathbf{x}^t$ $(\mathbf{x} \in \mathbb{R}^n)$. Per veronderstelling zijn de rijen van A lineair onafhankelijk. Omdat $\bar{\boldsymbol{\mu}}_L \neq \mathbf{0}$, volgt dat $\bar{\boldsymbol{\mu}}_L \times A \neq \mathbf{0}$. Zij $\mathbf{c} \in \mathbb{R}^n$ zó dat $\bar{\boldsymbol{\mu}}_L \times A \times \mathbf{c}^t > 0$. Alle positieve veelvouden van $\mathbf{c}$ voldoen ook aan deze strikte ongelijkheid. Dit impliceert dat er een $\mathbf{x}^\star \in \mathbb{R}^n$ is waarvoor

$$\sum_{i \in L} \bar{\mu}_i (d_i - g_i(\mathbf{x}^\star)) > 0.$$

Omdat elke g_i convex is. volgt dat deze ongelijkheid niet alleen voor $\mathbf{x}^\star$, maar ook voor elke $\mathbf{x}$ van de vorm

$$\mathbf{x} := (1 - \alpha)\mathbf{a} + \alpha \mathbf{x}^\star \ (0 < \alpha < 1)$$

geldt:

$$\sum_{i \in L} \bar{\mu}_i (d_i - g_i((1-\alpha)\mathbf{a} + \alpha \mathbf{x}^\star)) \geq \sum_{i \in L} \bar{\mu}_i (d_i - (1-\alpha) g_i(\mathbf{a}) - \alpha g_i(\mathbf{x}^\star))$$
$$= \sum_{i \in L} \bar{\mu}_i (d_i - (1-\alpha) d_i - \alpha g_i(\mathbf{x}^\star)) = \sum_{i \in L} \bar{\mu}_i (d_i - (1-\alpha) d_i - \alpha g_i(\mathbf{x}^\star))$$
$$= \alpha \sum_{i \in L} \bar{\mu}_i (d_i - g_i(\mathbf{x}^\star)) > 0.$$

Omdat $\mathbf{a} \in X$, $\mathbf{a}^I \gg \mathbf{0}$ en $\boldsymbol{\mu}^K \geq \mathbf{0}$, volgt dat er een $\mathbf{x} \in X'$ is waarvoor (3) niet geldt. Dit is een tegenspraak. Dus

$$\bar{v} > 0.$$

Zet nu

$$\boldsymbol{\lambda} := \boldsymbol{\mu}/\bar{v}.$$

We bewijzen nu dat $\boldsymbol{\lambda}$ als gezocht is. Allereerst merken we op dat $\boldsymbol{\lambda} \geq \mathbf{0}$, $\boldsymbol{\lambda}^K \geq \mathbf{0}$ en

$$\mathcal{L}(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}}).$$

Voor alle $\mathbf{x} \in X'$ volgt, vanwege (3) en (4),

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{j \in M} \lambda^j (d_j - g_j(\mathbf{x})) \leq f(\bar{\mathbf{x}}) = f(\bar{\mathbf{x}}) + \sum_{j \in M} \lambda^j (d_j - g_j(\bar{\mathbf{x}})) = \mathcal{L}(\bar{\mathbf{x}}, \boldsymbol{\lambda}).$$

En, voor alle $\boldsymbol{\eta} \in \mathbb{R}^m$ met $\boldsymbol{\eta}^K \geq 0$, geldt

$$\mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = f(\bar{\mathbf{x}}) \leq f(\bar{\mathbf{x}}) + \sum_{j \in M} \eta^j (d_j - g_j(\bar{\mathbf{x}})) = \mathcal{L}(\bar{\mathbf{x}}, \boldsymbol{\eta}). \quad \text{Q.E.D.}$$

Corollarium 2 *Beschouw het concave programmeringsprobleem ♠ waarbij aan de <u>conditie van Slater</u> is voldaan. Dan geldt voor alle* $\bar{\mathbf{x}} \in X$ *dat* $\bar{\mathbf{x}}$ *een maximaliseerder is d.e.s.d.a. er een* $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^m$ *met* $\bar{\boldsymbol{\lambda}}^K \geq \mathbf{0}$ *bestaat zodanig dat* $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}})$ *een zadelpunt van de lagrangefunctie* $\mathcal{L}$ *is.* ⋄

Als we de conditie van Slater in Propositie 3 weglaten, dan is het resultaat daar niet meer in het algemeen waar. Inderdaad: beschouw het concave programmeringsprobleem

$$\underset{\text{o.d.r. } x^2 \leq 0,\, x \geq 0}{\text{MAX}_{x \in \mathbb{R}}} \quad -x.$$

Aan de conditie van Slater is niet voldaan. Het is duidelijk dat 0 een maximaliseerder is. De lagrange-functie $\mathcal{L} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is $\mathcal{L}(x, \lambda) = x - \lambda x^2$. Er is echter geen $\bar{\lambda} \geq 0$ zodanig dat $(0, \bar{\lambda})$ een zadelpunt is. Immers aan $\mathcal{L}(x, \bar{\lambda}) \leq \mathcal{L}(0, \bar{\lambda}) \leq \mathcal{L}(0, \lambda)$ $(x \geq 0,\ \lambda \geq 0)$, i.e. aan $x - \bar{\lambda} x^2 \leq 0 \leq 0$ $(x \geq 0)$ voldoet geen enkele $\bar{\lambda} \geq 0$.

3 Differentieerbaarheid

In deze paragraaf beperken we ons verder tot niet-lineaire en concave programmeringsproblemen met louter ongelijkheidsresricties, i.e. tot $\spadesuit_{\leq}$, en halen we differentieerbaarheid erbij.[9]

Er bestaan ook varianten van de resultaten in deze paragraaf voor het niet-lineaire programmeringsprobleem. Zie daartoe bijvoorbeeld Jungnickel (1999).

3.1 Karakterisatie van zadelpunten

Propositie 4 *Beschouw het niet-lineaire programmeringsprobleem*

$$\text{MAX}_{\substack{\mathbf{x}\in U \\ o.d.r.\ \mathbf{g}(\mathbf{x})\leq\mathbf{0},\ \mathbf{x}\geq\mathbf{0}}} f(\mathbf{x}), \qquad \spadesuit_{\leq}$$

waar U een open deel van $\mathbb{R}^n$ is en $f, g_1, \ldots, g_m : U \to \mathbb{R}$ differentieerbare functies zijn.

Zij $\mathcal{L} : U \times \mathbb{R}^m \to \mathbb{R}$ de lagrange-functie geassocieerd met dat probleem, i.e.

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}).$$

Dan geldt voor alle $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \in (U \cap \mathbb{R}^n_+) \times \mathbb{R}^m_+$:

$$(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \textit{ is een zadelpunt van } \mathcal{L}$$

$\Rightarrow$

i. voor alle i: $\frac{\partial \mathcal{L}}{\partial x_i}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \leq 0$;

ii. voor alle i: $\frac{\partial \mathcal{L}}{\partial x_i}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = 0$ als $\bar{x}_i > 0$;

iii. voor alle j: $\frac{\partial \mathcal{L}}{\partial \lambda_j}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \geq 0$;

iv. voor alle j: $\frac{\partial \mathcal{L}}{\partial \lambda_j}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = 0$ als $\bar{\lambda}_j > 0$.

En als $f : U \to \mathbb{R}$ concaaf en $g_1, \ldots, g_m : U \to \mathbb{R}$ convex zijn, dan geldt hierboven ook "$\Leftarrow$". $\diamond$

Bewijs.— Stel dat $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}})$ een zadelpunt van $\mathcal{L}$ is. Volgens Propositie 2 is $\bar{\mathbf{x}}$ een maximaliseerder van het niet-lineaire programmeringsprobleem en $\mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = f(\bar{\mathbf{x}})$, dus

$$\sum_{i=1}^{m} \bar{\lambda}_i g_i(\bar{\mathbf{x}}) = 0.$$

Het eerste impliceert dat $g_j(\bar{\mathbf{x}}) \leq 0$ $(1 \leq j \leq m)$, i.e. dat iii geldt. En het tweede dat als $\bar{\lambda}_j > 0$ dat dan $g_j(\bar{\mathbf{x}}) = 0$ en dus dat iv geldt. Verder is er een $\delta > 0$ zodanig dat voor alle $\lambda \in [-\bar{x}_i, -\bar{x}_i + \delta)$

$$\mathcal{L}(\bar{\mathbf{x}} + \lambda \mathbf{e}_i, \bar{\boldsymbol{\lambda}}) \leq \mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}).$$

Dus $\lambda = 0$ is een maximaliseerder van de functie $\lambda \mapsto \mathcal{L}(\bar{\mathbf{x}} + \lambda \mathbf{e}_i, \bar{\boldsymbol{\lambda}})$ op $[-\bar{x}_i, -\bar{x}_i + \delta)$. Omdat deze functie differentieerbaar in 0 is, volgt

$$\frac{\partial \mathcal{L}}{\partial x_i}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \leq 0, \text{ en } \frac{\partial \mathcal{L}}{\partial x_i}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = 0 \text{ als } \bar{x}_i > 0.$$

Dus i en ii gelden.

[9] Maar zoals bijvoorbeeld in het reeds genoemde typoscript Balder (2003) na te lezen is, zijn resultaten als onderstaande ook te verkrijgen zonder differentieerbaarheid te veronderstellen. Dat komt omdat convexe functies weliswaar niet differentieerbaar hoeven te zijn, maar desalniettemin prettige differentieerbaarheidseigenschappen hebben. Het reddende object hier is dat van de subdifferentiaal.

Bewijzen we nu de omgekeerde bewering: stel dat $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \in (U \cap \mathbb{R}^n_+) \cap \mathbb{R}^m_+$ aan i - iv voldoet. De functie $U \to \mathbb{R}$ gedefinieerd door $\mathbf{x} \mapsto -\mathcal{L}(\mathbf{x}, \bar{\boldsymbol{\lambda}})$ is convex en differentieerbaar. Vanwege ☺ 6 steunt de affiene functie

$$T(\mathbf{x}) = -\mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) + \sum_{j=1}^{n} (x_j - \bar{x}_j) \frac{\partial - \mathcal{L}}{\partial x_j} (\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}})$$

deze functie te $\bar{\mathbf{x}}$. Omdat aan ii voldaan is geldt

$$T(\mathbf{x}) = -\mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) - \sum_{j=1}^{n} x_j \frac{\partial \mathcal{L}}{\partial x_j} (\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}).$$

Bovenstaande geeft voor alle $\mathbf{x} \in U \cap \mathbb{R}^n_+$ dat

$$\mathcal{L}(\mathbf{x}, \bar{\boldsymbol{\lambda}}) \leq -T(\mathbf{x}) \leq \mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}).$$

(Inderdaad, de eerste ongelijkheid geldt omdat T de functie $-\mathcal{L}(\cdot, \bar{\boldsymbol{\lambda}})$ steunt en de tweede volgt nu vanwege i.) Tenslotte volgt voor $\boldsymbol{\lambda} \geq \mathbf{0}$ met iv en iii dat

$$\mathcal{L}(\bar{\mathbf{x}}, \boldsymbol{\lambda}) = \mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) + \sum_{j=1}^{n} (\lambda_j - \bar{\lambda}_j) g_j(\bar{\mathbf{x}}) = \mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) + \sum_{j=1}^{n} \lambda_j g_j(\bar{\mathbf{x}}) \geq \mathcal{L}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}).$$

Dus $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}})$ is een zadelpunt van $\mathcal{L}$. Q.E.D.

Het bewijs van deze propositie heeft dus nog heel wat voeten in de aarde gehad: vooral de niet zo elementaire resultaten ☺ 6 en ☺ 7 uit de appendix hebben we gebruikt. Maar het bewijs dat we van de "$\Rightarrow$"-bewering leverden was recht-toe-recht-aan (en best slim).

Op dezelfde manier bewijst men:

Propositie 5 *Beschouw het niet-lineaire programmeringsprobleem*

$$\text{MAX}_{\substack{\mathbf{x} \in U \\ o.d.r.\ \mathbf{g}(\mathbf{x}) \leq \mathbf{0}}} f(\mathbf{x}), \quad \spadesuit_{\leq}$$

waar U een open deel van $\mathbb{R}^n$ is en $f, g_1, \ldots, g_m : U \to \mathbb{R}$ differentieerbare functies zijn.

Zij $\mathcal{L} : U \times \mathbb{R}^m \to \mathbb{R}$ de lagrange-functie geassocieerd met dat probleem, i.e.

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}).$$

Dan geldt voor alle $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \in U \times \mathbb{R}^m_+$:

$$(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \textit{ is een zadelpunt van } \mathcal{L}$$

$\Rightarrow$

i. voor alle i: $\frac{\partial \mathcal{L}}{\partial x_i}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = 0$;

ii. voor alle j: $\frac{\partial \mathcal{L}}{\partial \lambda_j}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) \geq 0$;

iii. voor alle j: $\frac{\partial \mathcal{L}}{\partial \lambda_j}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}) = 0$ als $\bar{\lambda}_j > 0$.

En als $f : U \to \mathbb{R}$ concaaf en $g_1, \ldots, g_m : U \to \mathbb{R}$ convex zijn, dan geldt hierboven ook "$\Leftarrow$". $\diamond$

3.2 De Stelling van Karush-Kuhn-Tucker

Stelling 2 *(Karush-Kuhn-Tucker.) Beschouw het concave programmeringsprobleem*

$$\underset{\substack{\mathbf{x}\in U \\ o.d.r.\ \mathbf{g}(\mathbf{x})\leq\mathbf{0},\ \mathbf{x}\geq\mathbf{0}}}{\mathrm{MAX}} f(\mathbf{x}),$$

waar U een open convex deel van $\mathbb{R}^n$ is en de functies $f, g_1, \ldots, g_m : U \to \mathbb{R}$ differentieerbaar zijn.

Stel er is een toelaatbaar element waar alle g_j negatief zijn. Met de functie $\mathcal{L} : U \times \mathbb{R}^m \to \mathbb{R}$ gedefinieerd door

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) := f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j g_j(\mathbf{x}),$$

geldt voor elke $\mathbf{x} \in U$ dat $\mathbf{x}$ een maximaliseerder is d.e.s.d.a. er een $\boldsymbol{\lambda} \in \mathbb{R}^m$ is zó dat

i. voor alle i: $x_i \geq 0$;

ii. voor alle j: $\lambda_j \geq 0$;

iii. voor alle i: $\frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}, \boldsymbol{\lambda}) \leq 0$ als $x_i = 0$;

iv. voor alle i: $\frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}, \boldsymbol{\lambda}) = 0$ als $x_i > 0$;

v. voor alle j: $\frac{\partial \mathcal{L}}{\partial \lambda_j}(\mathbf{x}, \boldsymbol{\lambda}) \geq 0$ als $\lambda_j = 0$;

vi. voor alle j: $\frac{\partial \mathcal{L}}{\partial \lambda_j}(\mathbf{x}, \boldsymbol{\lambda}) = 0$ als $\lambda_j > 0$. $\diamond$

Bewijs.— Zij $\mathbf{b}$ een toelaatbaar element waar alle g_i negatief zijn. Omdat U open is en de g_i en de functies x_i continu zijn, is er zelfs een toelaatbaar element $\mathbf{a}$ waar alle g_i negatief zijn met $\mathbf{a} \gg \mathbf{0}$. Dat betekent dat aan de conditie van Slater voldaan is.

Als $\mathbf{x}$ een maximaliseerder is, dan is er vanwege Corollarium 2 een $\boldsymbol{\lambda} \in \mathbb{R}^m_+$ zodanig dat $(\mathbf{x}, \boldsymbol{\lambda})$ een zadelpunt van $\mathcal{L}$ is. Pas nu Propositie 5 toe.

En als er een $\boldsymbol{\lambda} \in \mathbb{R}^n_+$ is zodanig dat aan i-vi voldaan is, dan is, volgens Propositie 5, $(\mathbf{x}, \boldsymbol{\lambda})$ een zadelpunt van $\mathcal{L}$ en dus, vanwege Corollarium 2, $\mathbf{x}$ een maximaliseerder. Q.E.D.

Naar de condities $i - vi$ verwijzen we ook wel als de karush-kuhn-tucker-condities (voor deze variant). Ze zijn equivalent met de volgende:

- voor alle i: $x_i \geq 0$;
- voor alle j: $\lambda_j \geq 0$;
- voor alle i: $\frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}, \boldsymbol{\lambda}) \leq 0$;
- voor alle i: $x_i \frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}, \boldsymbol{\lambda}) = 0$;
- voor alle j: $\lambda_j \frac{\partial \mathcal{L}}{\partial \lambda_j}(\mathbf{x}, \boldsymbol{\lambda}) = 0$.

Voorbeeld 1 Beschouw het maximalisatieprobleem

$$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2_+ \\ \text{o.d.r. } 3x_1+x_2\leq 150,\ x_1+x_2\leq 100}}{\mathrm{MAX}} x_1^{1/4} x_2^{1/2}.$$

Dit is een concaaf programmeringsprobleem van de vorm $\spadesuit_{\leq}$: neem $X = \mathbb{R}^2_+$, $f(x_1, x_2) = x_1^{1/4} x_2^{1/2}$, $g_1(x_1, x_2) = 3x_1 + x_2 - 150$ en $g_2(x_1, x_2) = x_1 + x_2 - 100$. (Merk nog onder andere op dat f inderdaad concaaf is.) Met $U = X$ is op deze situatie is de Stelling van Karush-Kuhn-Tucker niet toepasbaar omdat X niet open is. Ook is er geen open deelverzameling van $\mathbb{R}^2$ die $\mathbb{R}^2_+$ omvat waarop f wel-gedefinieerd is. Maar deze niet-toepasbaarheid kunnen we gemakkelijk verhelpen door op te merken dat $f = 0$ op $\mathbb{R}^2_+ \setminus \mathbb{R}^2_{++}$ en er een toegestaan element is waar f groter dan 0 is. Dit impliceer dat het maximalisatieprobleem

$$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2_{++} \\ \text{o.d.r. } 3x_1+x_2\le 150,\, x_1+x_2\le 100}}{\text{MAX}} \quad x_1^{1/4}x_2^{1/2}$$

dezelfde oplossing als het oorspronkelijke heeft en dat (nog opmerkend dat er een toelaatbaar element is waarin g_1 en g_2 negatief zijn) op dit nieuwe maximalisatieprobleem de Stelling van Karush-Kuhn Tucker wel toepasbaar is. Bekijken we verder daarom alleen dit nieuwe probleem. Omdat de doelfunctie nu strikt concaaf is, volgt uit ☺ 1, dat er ten hoogste één maximaliseerder is. Met ☺ 4 volgt nu zelfs dat er een unieke maximaliseerder is. De karush-kuhn-tucker-condities komen, met

$$\mathcal{L}(x_1, x_2) = x_1^{1/4}x_2^{1/2} - \lambda_1(3x_1 + x_2 - 150) - \lambda_2(x_1 + x_2 - 100)$$

neer op: $x_1, x_2 \in \mathbb{R}^2_{++}$ is een maximaliseerder d.e.s.d.a. er zijn $\lambda_1, \lambda_2 \geq 0$ zó dat $x_1, x_2, \lambda_1, \lambda_2$ een oplossing van het stelsel

$$\begin{aligned}
\frac{1}{4}x_1^{-3/4}x_2^{1/2} - 3\lambda_1 - \lambda_2 &= 0,\\
\frac{1}{2}x_1^{1/4}x_2^{-1/2} - \lambda_1 - \lambda_2 &= 0,\\
3x_1 + x_2 - 150 \leq 0 &\text{ als } \lambda_1 = 0,\\
x_1 + x_2 \leq 100 &\text{ als } \lambda_2 = 0,\\
3x_1 + x_2 - 150 = 0 &\text{ als } \lambda_1 > 0,\\
x_1 + x_2 = 100 &\text{ als } \lambda_2 > 0
\end{aligned}$$

is. Omdat we reeds weten dat er een unieke maximaliseerder is, volgt dat we zo gauw we een oplossing $x_1, x_2, \lambda_1, \lambda_2$ van bovenstaand stelsel vergelijkingen hebben, (x_1, x_2) die maximaliseerder is.

Laten we eens proberen of er een oplossing van het stelsel met $\lambda_1 = 0$ en $\lambda_2 = 0$ is. Wel, de eerste twee vergelijkingen laten zien dat zo'n oplossing er niet is.

Laten we nu eens proberen of $\lambda_1 > 0$ en $\lambda_2 > 0$ mogelijk is. In dat geval volgt $3x_1+x_2-150 = 0$ en $x_1 + x_2 = 100$ en daaruit $x_1 = 25, x_2 = 75$. Dit invullen in de eerste twee vergelijkingen van het stelsel lost deze twee op met $\lambda_1 = \frac{\sqrt{15}}{120}$ en $\lambda_2 = \frac{\sqrt{15}}{40}$.

Conclusie: $(25, 75)$ is de unieke maximaliseerder en $25^{1/4}\sqrt{75}$ is het maximum. ◇

4 Opgaven

Opgave 1 *Voer voor elk der volgende maximalisatieproblemen het volgende uit:*

1. *Bewijs dat het om een concaaf programmeringsprobleem van de vorm* $♠_{\leq}$ *gaat.*
2. *Bewijs dat er een unieke maximaliseerder is.*
3. *Laat zien dat Stelling 2 toepasbaar is en schrijf de karush-kuhn-tucker-condities op.*
4. *Bepaal het maximum en de unieke maximaliseerder*

a. $$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2 \\ o.d.r.\ x_1\le 3,\ x_2\le 2,\ x_1\ge 0, x_2\ge 0}}{\text{MAX}} \quad -(x_1 + 1)^2 - (x_2 - 5)^2.$$

b. $$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2 \\ o.d.r.\ x_1+x_2\le 2,\ x_1\ge 0, x_2\ge 0}}{\text{MAX}} \quad 6x_1 - 2x_1^2 + 2x_1x_2 - 2x_2^2.$$

c. $$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2 \\ o.d.r.\ x_1+x_2\le 1,\ x_1\ge 0, x_2\ge 0.}}{\text{MAX}} \quad x_1 + 2x_2 + 2x_1x_2 - \frac{x_1^2}{2} - \frac{x_2^2}{2}.$$

Opgave 2 *Beschouw het maximalisatieprobleem*

$$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2 \\ o.d.r.\ x_1^2-x_2\leq 0,\ x_1+x_2-6\leq 0,\ -x_1+1\leq 0,\ x_1\geq 0,\ x_2\geq 0}}{\text{MAX}} \quad -(x_1-\frac{9}{4})^2-(x_2-2)^2.$$

Los dit probleem grafisch op en ook met de in het typoscript ontwikkelde theorie.

Opgave 3 *Los het volgene maximalisatieprobleem op:*

$$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2 \\ o.d.r.\ 4-(x_1+1)(x_2+1)\geq 0,\ x_1\geq 0, x_2\geq 0}}{\text{MAX}} \quad 2x_1+x_2.$$

Opgave 4 *Beschouw het maximalisatieprobleem*

$$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2 \\ o.d.r.\ x_1+x_2\leq 6,\ x_1\geq 0,\ x_2\geq 0}}{\text{MAX}} \quad 2x_1+3x_2.$$

Los dit probleem eerst grafisch op en daarna ook formeel.

Opgave 5 *Los het volgende minimalisatieprobleem op:*

$$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2 \\ o.d.r.\ x_1^2+x_2^2\leq 5, x_1+x_2\leq 3,\ x_1\geq 0, x_2\geq 0}}{\text{MIN}} \quad (x_1-3)^2+(x_2-2)^2.$$

Opgave 6 *Los de volgende maximalisatieprobleem op:*

a. $\underset{\substack{(x_1,x_2)\in\mathbb{R}^2 \\ o.d.r.\ x_1^2+x_2^2\leq 1,\ x_1\geq 0,\ x_2\geq 0}}{\text{MAX}} \quad x_1+ax_2,\ waar\ a\geq 0.$

b. $\underset{\substack{(x_1,x_2)\in\mathbb{R}^2 \\ o.d.r.\ x_1^2+x_2^2\leq 1,\ x_1+x_2\geq 0}}{\text{MAX}} \quad x_1+ax_2,\ waar\ a\in\mathbb{R}\ is.$

Opgave 7 *Los het volgende maximalisatieprobleem op:*

$$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2_+ \\ o.d.r.\ 3x_1+x_2\leq 150,\ x_1+x_2\leq 100}}{\text{MAX}} \quad x_1^a x_2^b,$$

waar $a, b > 0$ *met* $a + b < 1$ *en* $a < b < 3a$.

Opgave 8 *Los het volgende maximalisatieprobleem op:*

$$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2_+ \\ o.d.r.\ 3x_1+x_2\leq 150,\ x_1+x_2\leq 100}}{\text{MAX}} \quad x_1^{1/2} x_2^{1/4}.$$

Opgave 9 *Los het volgende maximalisatieprobleem op:*

$$\underset{\substack{(x_1,x_2)\in\mathbb{R}^2_+ \\ o.d.r.\ 3x_1+x_2\leq 150,\ x_1+x_2\leq 100,\ x_1\geq 30,\ x_1\leq 80}}{\text{MAX}} \quad x_1^{1/4} x_2^{1/2}.$$

A Appendix

A.1 Existentie en uniciteit van maximaliseerders

In deze deelparagraaf geven we we enkele elementaire resultaten af betreffende maximaliseerders voor het maximalisatieprobleem

$$\text{MAX}\, f,$$

waar f een functie op een verzameling X is.

☺ **1** *Indien X een convex deel van een lineaire ruimte is en f strikt quasi-concaaf is, dan is er hoogstens één maximaliseerder.*

Bewijs.— Als a, b twee verschillende maximaliseerders zouden zijn, dan zou met $f(a) = f(b) =: m$ de tegenspraak $f(\frac{1}{3}a + \frac{2}{3}b) > \frac{1}{3}f(a) + \frac{2}{3}f(b) = m$ gelden. Q.E.D.

☺ **2** *Indien X een convex deel van een lineaire ruimte is en f quasi-concaaf is, dan is de verzameling der maximaliseerders convex.*

Bewijs.— Als a, b twee maximaliseerders zijn en $\lambda, \mu \geq 0$ zijn met som 1, dan geldt voor alle $x \in X$ dat $f(\lambda a + \mu b) \geq \min(f(a), f(b)) = f(a) \geq f(x)$. Dus ook $\lambda a + \mu b$ is een maximaliseerder. Q.E.D.

☺ **3** *Indien f continu is en X een metrische ruimte is, dan is de verzameling der maximaliseerders van f gesloten.*

Bewijs.— Stel (a_n) is een convergente rij van maximaliseerders van f met limiet a. Dan $f(a_n) = f(a_1)$ voor alle n. Omdat f continu is volgt $f(a) = f(\lim_{n\to\infty} a_n) = \lim_{n\to\infty} f(a_n) = f(a_1)$. Dus ook a is een maximaliseerder van f. Q.E.D.

☺ **4** *(Stelling van Weierstrass.) Indien X een niet-lege compacte metrische ruimte is en f continu is, dan heeft f een maximaliseerder en is de verzameling der maximaliseerders van f compact.*

Bewijs.— Omdat f continu is en X compact is, is $f(X)$ is een compact deel van $\mathbb{R}$ en vandaar gesloten en begrensd. Dat impliceert dat het supremum m van $f(X)$ een reëel getal is (i.e. niet $+\infty$ is). Er geldt $m \in \overline{f(X)} = f(X)$. Dus is er een $a \in X$ met $m = f(a)$ en daarmee is a een maximaliseerder. Vanwege ☺ 3 is de verzameling der maximaliseerders van f gesloten. Omdat X compact is, is deze verzameling compact. Q.E.D.

☺ **5** *Indien X een topologische ruimte is, A een deelverzameling van X met afsluiting gelijk aan X is en f continu is, dan is elke maximaliseerder van $f \restriction A$ ook een maximaliseerder van f.*

Bewijs.— Uit het ongerijmde. Stel $x \in A$ is een maximaliseerder van $f \restriction A$ maar niet van f. Dan is er een $y \in X$ met $f(y) > f(x)$. Natuurlijk is $y \notin A$. Omdat $X = \bar{A} = A \cup \partial A$, is $y \in \partial A$. Omdat f continu te y is, is er een $\delta > 0$ zodanig dat $f > f(x)$ op $B_\delta(y)$. Omdat $y \in \partial A$ bevat $B_\delta(y)$ een element van A, zeg z. Omdat $f(z) > f(x)$ hebben we een tegenspraak. Q.E.D.

A.2 Convex gereedschap

☺ **6** *Zij A een open convex deel van $\mathbb{R}^n$ en $f : A \to \mathbb{R}$ een convexe functie. Laat $\mathbf{a} \in A$. Dan: f is differentieerbaar in $\mathbf{a}$ $\Leftrightarrow$ er is een unieke affiene functie $T : \mathbb{R}^n \to \mathbb{R}$ die f te $\mathbf{a}$ steunt; d.w.z. $T(\mathbf{a}) = f(\mathbf{a})$ en $T(\mathbf{a}) \leq f(\mathbf{a})$. En in dat geval is $T(\mathbf{x}) := f(\mathbf{a}) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{a})(x_j - a_j)$ die functie.*

☺ **7** *Zij D een convex deel van $\mathbb{R}^n$, $f_1, \ldots, f_p : D \to \mathbb{R}$ concave functies, $g_1, \ldots, g_m : D \to \mathbb{R}$ concave functies en $h_1, \ldots, h_l : D \to \mathbb{R}$ affiene functies. Stel*

$$\{\mathbf{x} \in D \mid f_j(\mathbf{x}) \leq 0 \ (1 \leq j \leq p)\} \supseteq \{\mathbf{x} \in D \mid g_i(\mathbf{x}) \geq 0 \ (1 \leq i \leq m) \text{ en } h_k(\mathbf{x}) = 0 \ (1 \leq k \leq l)\}.$$

Dan

1. *Er bestaan $v_1, \ldots, v_p, u_1, \ldots, u_m \geq 0$ en $w_1, \ldots, w_k \in \mathbb{R}$ maar niet allen gelijk aan 0, zodanig dat*
$$\sum_{j=1}^{p} v_j f_j(\mathbf{x}) + \sum_{i=1}^{m} u_i g_i(\mathbf{x}) + \sum_{k=1}^{l} w_k h_k(\mathbf{x}) \leq 0 \ (\mathbf{x} \in D).$$

2. *Als er een $\mathbf{x} \in D$ bestaat met $g_i(\mathbf{x}) \geq 0$ $(1 \leq i \leq m)$, $f^j(\mathbf{x}) = 0$ $(1 \leq j \leq p)$ en $h_k(\mathbf{x}) = 0$ $(1 \leq k \leq l)$, dan geldt $\sum_{i=1}^{m} u_i g_i(\mathbf{x}) = 0$.*

☺ **8** *Zij* X *een convex deel van* $\mathbb{R}^n$ *en* $g : X \to \mathbb{R}$ *een functie.* g *heet* *<u>pseudo-convex</u>* *in een inwendig punt* $\mathbf{y}$ *van* X *als* g *in* $\mathbf{y}$ *differentieerbaar is en voor alle* $\mathbf{x} \in X$ *de implicatie*

$$Dg(\mathbf{y})(\mathbf{x} - \mathbf{y}) \geq 0 \;\Rightarrow g(\mathbf{x}) \geq g(\mathbf{y})$$

geldt. Als X *open is en* g *in elk punt van* X *pseudo-convex is, dan heet* g *<u>pseudo-convex</u>. Dus een functie* g *op een open convex deel* X *van* $\mathbb{R}^n$ *is pseudo-convex d.e.s.d.a. voor alle* $\mathbf{x}, \mathbf{y} \in X$ *er geldt* $Dg(\mathbf{y})(\mathbf{x} - \mathbf{y}) \geq 0 \;\Rightarrow g(\mathbf{x}) \geq g(\mathbf{y})$.

g *heet* *<u>pseudo-concaaf</u>* *(in* $\mathbf{y}$*) als* $-g$ *(in* $\mathbf{y}$*) pseudo-convex is.*

☺ **9** *(Convexiteitsnoties vergeleken.) Zij* X *een convex deel van* $\mathbb{R}^n$ *en* $f : X \to \mathbb{R}$ *een functie.*

1. *Stel* f *is convex. Als* f *in een inwendig punt* $\mathbf{y}$ *van* X *differentieerbaar is, dan is* f *pseudo-convex in* $\mathbf{y}$*. Als* X *tevens open is en* f *differentieerbaar is, dan is* f *pseudo-convex.*

2. *Zij* f *continu en quasi-convex. Dan is* f *in elk inwendig punt van* X *waar* f *differentieerbaar is met afgeleide ongelijk nul pseudo-convex.*

3. *Stel* X *is open en* f *is differentieerbaar. Als* f *pseudo-convex is, dan is* f *quasi-convex. En als de afgeleide van* f *nergens nul is, dan geldt:* f *is quasi-convex* $\Leftrightarrow$ f *is pseudo-convex.*

Referenties

E. Balder. Working with the Kuhn-Tucker condtions. Technical report, Mathematisch Instituut Utrecht, Utrecht, 2003.

J. Bracken and P. McCormick. *Selected Applications of Nonlinear Programming.* Wiley, New York, 1968.

D. Jungnickel. *Optimierungsmethoden.* Springer-Verlag, Berlin, 1999. ISBN 3 540 66057 7.

M. Truchon. *Théorie de l'Optimisation Statique et Différentiable.* Gaëtan morin, Cicoutimi, 1987. ISBN 2 89105 118 1.