

Zeitschrift: Elemente der Mathematik
Herausgeber: Schweizerische Mathematische Gesellschaft
Band: 18 (1963)
Heft: 1

Artikel: Nichtlineare Programmierung
Autor: Künzli, H.P. / Oettli, W.
DOI: <https://doi.org/10.5169/seals-22631>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 10.04.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

ELEMENTE DER MATHEMATIK

Revue de mathématiques élémentaires – Rivista di matematica elementare

*Zeitschrift zur Pflege der Mathematik
und zur Förderung des mathematisch-physikalischen Unterrichts
Organ für den Verein Schweizerischer Mathematik- und Physiklehrer*

Publiziert mit Unterstützung des Schweizerischen Nationalfonds
zur Förderung der wissenschaftlichen Forschung

El. Math. Band XVIII Nr. 1 Seiten 1–24 Basel, 10. Januar 1963

Nichtlineare Programmierung*)

Herrn Prof. Dr. B. L. VAN DER WAERDEN zum 60. Geburtstag (2. Februar 1963) gewidmet

Unter mathematischer Programmierung versteht man die Anwendung mathematischer Methoden auf Planungsprobleme der Unternehmensforschung, wie sie etwa bei der Preis-, Produktions- und Reklameplanung einer Firma auftreten. Man unterscheidet im wesentlichen zwischen der linearen Programmierung mit ihren Erweiterungen wie ganzzahliger, parametrischer, stochastischer, dynamischer Programmierung und der nichtlinearen Programmierung.

Bei einem linearen Programm handelt es sich bekanntlich um die folgende Aufgaben:¹⁾

$$\text{I} \left\{ \begin{array}{l} \text{Die lineare (Ziel-) Funktion} \\ \sum_{i=1}^n p_i x_i \quad (1) \\ \text{ist zum Maximum zu machen unter den Nebenbedingungen} \\ \sum_{i=1}^n a_{ji} x_i \leq b_j \quad (j = 1, 2, \dots, m), \quad (2) \\ x_i \geq 0 \quad (i = 1, 2, \dots, n). \quad (3) \end{array} \right.$$

Um die Restriktionen (2) in Gleichungsform zu bringen, führt man m nichtnegative «Schlupfvariable» y_j ein und ersetzt (2) durch

$$\sum_{i=1}^n a_{ji} x_i + y_j, \quad y_j \geq 0 \quad (j = 1, 2, \dots, m). \quad (2')$$

In unmittelbar verständlicher Matrix-Vektorschreibweise ist Problem I dann äquivalent dem folgenden Problem:

$$\text{II} \left\{ \begin{array}{l} \text{Zu maximieren ist } p'x \text{ unter den Nebenbedingungen} \\ Ax + y = b, \quad (4) \\ x \geq 0, \quad y \geq 0. \quad (5) \end{array} \right.$$

*) Die vorliegende Darstellung entspricht einer kurzen Zusammenfassung des kürzlich erschienenen Werkes über Nichtlineare Programmierung von Künzi und Krelle (vgl. Literaturverzeichnis).

¹⁾ Vgl. hierzu auch den früheren Artikel über Lineare Programmierung (El. Math. XVI, 1-12 (1961)).

(nach Abzug der Stückkosten) die Nettopreise p_1, \dots, p_n je Einheit lösen kann. Für die Produktion benötigt man m nicht in beliebiger Menge vorhandene Produktionsfaktoren, und zwar pro Einheit des i -ten Produktes die Menge a_{ji} des Produktionsfaktors j . Die Produktionsfaktoren stehen nur bis zu den Höchstmengen b_1, \dots, b_m zur Verfügung. Gesucht wird das optimale Produktionsprogramm. Diese Aufgabe wird offensichtlich durch das Programm I gelöst.

In jüngster Zeit hat die nichtlineare Programmierung wachsende Bedeutung erlangt, da sie häufig wirklichkeitsnähere Lösungen ermöglicht. Wenn beispielsweise die «Preise» p_i nicht mehr konstant sind, sondern einem von x abhängigen Nachfragegesetz gehorchen, so hört die Zielfunktion (1) auf, linear zu sein.

Das allgemeine nichtlineare Programm würde etwa darin bestehen, eine Zielfunktion

$$\text{IV} \left\{ \begin{array}{l} F(x) \\ \text{zu minimieren unter den Nebenbedingungen} \\ f_j(x) \leq 0 \quad (j = 1, \dots, m), \\ x \geq 0. \end{array} \right. \quad \begin{array}{l} (9) \\ (10) \\ (11) \end{array}$$

Man muss jedoch vorerst noch die Einschränkung treffen, dass sowohl $F(x)$ wie $f_j(x)$ konvexe Funktionen sind. Damit ist sichergestellt, dass der durch die Nebenbedingungen gegebene zulässige Bereich konvex ist, und dass jedes lokale Minimum der Zielfunktion in diesem Bereich, das durch differentielle Kriterien gefunden werden kann, zugleich das gesuchte globale Minimum darstellt.

Das Theorem von KUHN und TUCKER, das eine zentrale Rolle in der konvexen Programmierung spielt, gibt notwendige und hinreichende Bedingungen für die Optimallösung von IV, und zwar mittels einer verallgemeinerten Lagrangefunktion L , die unter Verwendung von m Multiplikatoren u_1, \dots, u_m , kurz u , nach der Vorschrift

$$L(x, u) = F(x) + \sum_{j=1}^m u_j f_j(x) \quad (12)$$

gebildet wird. Danach stellt ein Vektor x dann und nur dann eine Lösung von IV dar, wenn er zusammen mit einem Vektor u folgende Bedingungen erfüllt:

$$L_x \geq 0; \quad x \geq 0; \quad x' L_x = 0; \quad (13)$$

$$L_u \leq 0; \quad u \geq 0; \quad u' L_u = 0. \quad (14)$$

Hierbei bedeutet L_x den Spaltenvektor der partiellen Ableitungen von L nach den x_i , genommen an der Stelle (x, u) , analog für L_u . Substituiert man aus (12), so besagt $L_u \leq 0$ zusammen mit $x \geq 0$ nur, dass die Lösung im zulässigen Bereich liegen muss. Die übrigen Bedingungen dagegen charakterisieren die Lösung wesentlich. Sie besagen anschaulich, dass im Optimalpunkt x der negative Gradientenvektor der Zielfunktion (der in Richtung des steilsten Abfalles von $F(x)$ weist) darstellbar sein muss als nichtnegative Linearkombination der nach aussen weisenden Normalen derjenigen berandenden Hyperflächen $f_j(x) = 0$ bzw. $x_i = 0$, auf denen x liegt.

Wenn in IV die Restriktionen (11) fehlen oder bereits unter (10) aufgenommen sind, so dass die x_i , formal wenigstens, nicht vorzeichenbeschränkt sind, so wird die Bedingung (13) ersetzt durch

$$L_x = 0. \quad (15)$$

ZOUTENDIJK, ROSEN, KELLEY und andere haben verschiedene Verfahren vorgeschlagen, die es erlauben, die Lösung eines allgemeinen konvexen Programmes beliebig genau anzunähern. Da diese Verfahren jedoch alle einen ziemlichen Aufwand erfordern, beschränken wir uns hier auf den Spezialfall quadratischer Zielfunktion und linearer Restriktionen, für den exakte Lösungsverfahren vorliegen. Es sei also $F(x) = p'x + \frac{1}{2} x' C x = Q(x)$ eine quadratische Funktion. C ist eine n -reihige symmetrische Matrix, die wegen der geforderten Konvexität von $F(x)$ positiv semi-definit sein muss. Ferner sei $f_j(x) = a'_j x - b_j$. Fassen wir die m Zeilenvektoren α'_j zu einer $(m \cdot n)$ -Matrix A zusammen, so besteht das quadratische Programm, ein Sonderfall von IV, darin,

$$V \quad \left\{ \begin{array}{l} Q(x) = p'x + \frac{1}{2} x' C x \end{array} \right. \quad (16)$$

zu minimieren unter den Nebenbedingungen

$$A x \leq b, \quad (17)$$

$$x \geq 0. \quad (18)$$

Die Lagrangefunktion für dieses Problem lautet

$$L(x, u) = Q(x) + \sum_{j=1}^m u_j (a'_j x - b_j) = p'x + \frac{1}{2} x' C x + u' (A x - b). \quad (19)$$

Setzen wir zur Abkürzung $L_x = v$, $L_u = -y$, so ist offenbar $v = p + Cx + A'u$, $y = -Ax + b$. Die Bedingungen für die optimale Lösung lauten dann:

$$VI \quad \left\{ \begin{array}{l} A x \quad \quad \quad + y = b, \end{array} \right. \quad (20)$$

$$C x - v + A' u \quad = -p, \quad (21)$$

$$x \geq 0, \quad v \geq 0, \quad u \geq 0, \quad y \geq 0. \quad (22)$$

und

$$x' v + y' u = 0. \quad (23)$$

Dabei entsprechen (22) und (23) den eigentlichen Kuhn-Tucker-Bedingungen (13), (14), während (20) und (21) den Definitionsgleichungen für y und v entsprechen. Ein n -Vektor x stellt also dann und nur dann eine Lösung für V dar, wenn er zusammen mit einem n -Vektor v und m -Vektoren u , y den obigen Bedingungen geneigt. Diese Schreibweise der Kuhn-Tucker-Bedingungen für den quadratischen Fall geht auf BARANKIN und DORFMAN zurück.

Die Bedingungen VI bilden ein lineares System. Bedingung (23) lässt sich kombinatorisch so aussprechen, dass von je zwei nichtnegativen Variablen x_i und v_i bzw. u_j und y_j mindestens eine Null sein muss. Es kommen somit als Anwarter darauf, auch

(23) zu genügen, nur solche Lösungen des Systems VI in Frage, bei denen höchstens soviele Variable positiv sind, wie VI Gleichungen aufweist, nämlich $m + n$; das sind aber gerade die zulässigen Basislösungen von VI. Um unter den zulässigen Basislösungen von VI eine ausfindig zu machen, die auch (23) genügt, liegt der Versuch nahe, in abgewandelter Form auf das Simplexverfahren der linearen Programmierung zurückzugreifen, da sich dieses mit den Basislösungen eines linearen Systems beschäftigt. Das Verfahren von WOLFE zeigt eine Möglichkeit hierfür. Es besteht grob darin, dass man in (20) und (21) zunächst Abweichungen in Form künstlicher Schlupfvariablen zulässt und diese dann mittels des Simplexverfahrens zu Null reduziert, jedoch so, dass (23) im ganzen Verlauf des Verfahrens erfüllt ist. Man geht so vor, dass man VI ersetzt durch das erweiterte System

$$\text{VII} \quad \left\{ \begin{array}{l} Ax + y + D_1 w = b, \\ Cx - v + A'u + D_2 z = -p, \\ x \geq 0, \quad v \geq 0, \quad u \geq 0, \quad y \geq 0, \quad w \geq 0, \quad z \geq 0. \end{array} \right.$$

Hierbei sind D_1 und D_2 m - bzw. n -reihige Diagonalmatrizen mit den Elementen $+1$ oder -1 , je nachdem ob die rechte Seite der entsprechenden Gleichung positiv oder negativ ist. Für VII kann man sofort eine zulässige Basislösung angeben, nämlich diejenige mit $x = 0$, $y = 0$, $v = 0$, $u = 0$. Man verwendet die Simplexmethode, um, ausgehend von dieser Basislösung, die Linearform

$$\sum_{j=1}^m w_j$$

zu minimieren unter den Nebenbedingungen VII. Die v_i und u_j dürfen dabei nicht in die Basis genommen werden; sie behalten den Wert Null. Als Resultat erhält man eine Basislösung von VII mit $w = 0$; alle w_j sind aus der Basis verschwunden.

In einer zweiten Phase minimiert man, ausgehend von der am Schluss der ersten Phase erhaltenen Basislösung, die Linearform

$$\sum_{i=1}^n z_i$$

unter den Nebenbedingungen VII und mit der Zusatzvorschrift, dass v_i nicht in die Basis genommen werden darf, wenn x_i sich bereits in der Basis befindet, und umgekehrt; ebenso für u_j und y_j . Damit ist sichergestellt, dass für alle auftretenden Basislösungen Bedingung (23) erfüllt ist. Die w_j behalten in der zweiten Phase den Wert Null; sie dürfen also auch nicht in die Basis zurückkehren.

Wenn das Minimum, das man auf diese Weise für $\sum z_i$ erreicht, den Wert Null hat, so hat man am Ende der zweiten Phase eine Basislösung von VII mit $w = 0$, $z = 0$ und damit auch eine Lösung von VI erhalten. Da überdies (23) erfüllt ist, ist man am Ziel; der x -Teil löst das Programm V. Dieser Fall tritt aber nur mit Sicherheit ein, wenn entweder $p = 0$ oder C streng definit ist. Ist diese Voraussetzung nicht erfüllt, so benötigt man noch eine dritte Phase. Man ersetzt in den ersten beiden Phasen den Vektor p in VII durch $u \cdot p$, wobei u eine zusätzliche Variable ist, die nicht in die Basis genommen werden darf, sondern den Wert Null behält. Das entspricht $p = 0$,

und die zweite Phase endet infolgedessen mit $z = 0$. In der dritten Phase maximiert man die triviale Linearform u , wieder unter der Zusatzvorschrift und mit $w = 0$, $z = 0$. Es ergibt sich eine Folge von Basislösungen mit monoton wachsendem u . Man wählt eine konvexe Linearkombination zweier aufeinanderfolgender Basislösungen derart, dass die gleiche Linearkombination der zugehörigen u -Werte eins ergibt; der x -Teil der Linearkombination löst V.

Einen anderen Weg, den Bedingungen VI und (23) mittels der Simplextechnik beizukommen, zeigt das Verfahren von BARANKIN und DORFMAN. Während WOLFE (23) immer erfüllt sein lässt und dafür zunächst Abweichungen vom Hauptsystem VI in Kauf nimmt, sind bei BARANKIN und DORFMAN die Bedingungen VI immer streng erfüllt. Man geht aus von einer zulässigen Basislösung von VI, die nicht notwendig (23) erfüllt, und reduziert über eine Folge von Basislösungen die nichtlineare Zielfunktion $T = x'v + y'u$ zu Null. Die Transformationsformeln für den Übergang von einer Basislösung zur nächsten sind die gleichen wie beim Simplexverfahren der linearen Programmierung. Lediglich die Auswahlregel für die in die Basis zu nehmende Variable ist wegen der Nichtlinearität von T komplizierter. Man kann allerdings Beispiele konstruieren, bei denen das Verfahren von BARANKIN und DORFMAN steckenbleibt, weil sich keine Simplextransformation mehr durchführen lässt derart, dass T abnimmt, obwohl immer noch $T > 0$. FRANK und WOLFE haben diese Schwierigkeit beseitigt, indem sie T geeignet linearisierten.

Ein Verfahren von BEALE verwendet den Simplexmechanismus direkt, ohne den Umweg über VI. Man startet mit einer zulässigen Basislösung des Systems $Ax + y = b$, $x \geq 0$, $y \geq 0$. Die $m + n$ Variablen x und y seien zu einem Vektor z zusammengefasst. Mittels der Gleichungsrestriktionen lassen sich die m positiven Basisvariablen, etwa z_{n+1} bis z_{n+m} , sowie Q darstellen als Funktion der unabhängigen, in der Basislösung verschwindenden Variablen z_1 bis z_n :

$$z_g = d_{g_0} + \sum_{h=1}^n d_{gh} z_h \quad (g = n+1, \dots, n+m),$$

$$Q = c_{00} + c_{01} z_1 + \dots + c_{0n} z_n$$

$$+ (c_{10} + c_{11} z_1 + \dots + c_{1n} z_n) \cdot z_1$$

$$\vdots$$

$$+ (c_{n0} + c_{n1} z_1 + \dots + c_{nn} z_n) \cdot z_n \quad (c_{ij} = c_{ji}).$$

Die Klammerausdrücke entsprechen den partiellen Ableitungen $\frac{1}{2} \partial Q / \partial z_h$.

Falls alle Ableitungen ≥ 0 sind, so ist die Basislösung optimal. Andernfalls erhält man eine der Variablen mit negativer Ableitung, etwa z_1 , und zwar solange, bis eine der abhängigen Variablen oder $\frac{1}{2} \partial Q / \partial z_1$ verschwindet. Im ersten Falle ersetzt die neue verschwindende Variable z_1 als unabhängige Variable; z_1 ist basisch geworden. Die zu dem neuen Satz von unabhängigen Variablen gehörende Basislösung liefert einen geringeren Q -Wert. Im zweiten Fall führt man eine freie, das heisst nicht vorzeichenbeschränkte Variable u_1 ein, die mit den bisherigen unabhängigen Variablen gemäss

$$u_1 = \frac{1}{2} \frac{\partial Q}{\partial z_1} = c_{10} + \sum_{h=1}^n c_{1h} z_h$$

verknüpft ist. u_1 ersetzt z_1 als unabhängige Variable; z_1 wird abhängig von u_1 und z_2 bis z_n . Setzt man den neuen Satz von unabhängigen Variablen gleich Null, so hat man wieder eine Basislösung; allerdings hat sich die Anzahl der Basisvariablen um eins erhöht, da z_1 in die Basis kam, ohne dass eine der bisherigen Basisvariablen zum Verschwinden gebracht worden wäre. In beiden Fällen kann man zur Transformation der d_{gh} und, bei geeignetem Vorgehen, auch zur Transformation der c_{ij} auf die neue Basislösung die Formeln des Simplexverfahrens verwenden.

Wenn sich unter den unabhängigen Variablen freie Variable befinden, so müssen am Optimalpunkt die Ableitungen nach den freien Variablen verschwinden. Wenn die Ableitung nach einer freien Variablen positiv ist, so kann man Q noch erniedrigen, indem man die freie Variable negativ werden lässt, was bei den eigentlichen Variablen z_h wegen der Vorzeichenbeschränkung nicht möglich ist. Wenn man noch die zusätzliche Vorschrift einführt, dass wenn immer möglich eine freie unabhängige Variable in die Basis genommen, das heisst positiv oder negativ gemacht werden soll, so lässt sich zeigen, dass man nach endlich vielen Schritten zu einem Satz von unabhängigen Variablen gelangt, bei dem die Ableitungen von Q nach den z -Variablen nichtnegativ und die nach den freien Variablen null sind. Damit hat man das Optimum erreicht.

Entsprechend zu den dualen linearen Programmen kann man auch ein quadratisches Programm dualisieren, allerdings nicht mit der gleichen Symmetrie wie im linearen Fall. Einem gegebenen quadratischen Primärprogramm,

$$\text{VIII} \left\{ \begin{array}{l} p' x + \frac{1}{2} x' C x \\ \text{zu minimieren unter den Nebenbedingungen} \\ A x \leq b \end{array} \right. \quad \begin{array}{l} (16) \\ (17) \end{array}$$

wird zugeordnet das Dualprogramm,

$$\text{IX} \left\{ \begin{array}{l} -\frac{1}{2} x' C x - b' u \\ \text{zu maximieren unter den Nebenbedingungen} \\ C x + A' u = -p, \\ u \geq 0. \end{array} \right. \quad \begin{array}{l} (24) \\ (25) \\ (26) \end{array}$$

Es gelten die gleichen Sätze wie im linearen Fall: Wenn das eine Problem eine Lösung aufweist, so auch das andere, und die beiden Extremalwerte sind gleich. Wenn beim einen Problem die Zielfunktion über dem zulässigen Bereich nicht beschränkt ist, so ist beim anderen Problem der zulässige Bereich leer.

Wenn C streng definit und somit nichtsingulär ist, so kann man mittels der Gleichungsrestriktionen x aus dem dualen Programm eliminieren:

$$x(u) = -C^{-1} (A' u + p).$$

IX lautet dann:

$$X \quad \left\{ \begin{array}{l} -h' u - u' G u \\ \text{ist zu maximieren unter den Nebenbedingungen} \\ u \geq 0. \end{array} \right. \quad \begin{array}{l} (24') \\ \\ (26) \end{array}$$

Dabei ist zur Abkürzung $h = A C^{-1} p + b$ und $G = \frac{1}{2} A C^{-1} A'$ gesetzt. x stellt dann und nur dann eine Lösung von VIII dar, wenn $x = x(u)$, wobei u eine Lösung von X darstellt.

H. P. KÜNZI und W. OETTLI, Zürich

LITERATURVERZEICHNIS

- KRELLE, W., KUENZI, H. P., *Lineare Programmierung* (Industrielle Organisation, Zürich 1958).
 KUENZI, H. P., KRELLE, W., *Nichtlineare Programmierung*. (Springer-Verlag, Berlin-Göttingen-Heidelberg, 1962).
 VAJDA, S., *Mathematical Programming* (Addison-Wesley, London 1961).
 ZOUTENDIJK, G., *Methods of feasible Directions* (Elsevier, Amsterdam 1960).

Arithmetical Notes, XIII. A Sequel to Note IV*)

1. Introduction

First we introduce some terminology and notation. For any positive integer n ,

$$n = p_1^{e_1} \cdots p_r^{e_r}, \quad (1.1)$$

where p_1, \dots, p_r are the prime divisors of n , let $t(n)$ denote the maximal multiplicity of n , that is, $t(n) = \max(e_1, \dots, e_r)$, $t(1) = 0$. For arbitrary positive integers k , let n_k denote the k -segment of n , meaning the number $n_k = n/m$ where m is the largest k -th power divisor of n . Further, let $t_k(n)$ denote the maximal multiplicity of the k -segment of n , $t_k(n) = t(n_k)$.

In this note we prove elementary estimates for the distribution of two sequences of integers, $U_{k,t}$ and $V_{k,t}$, defined as follows. Let t denote a fixed positive integer; define $U_{k,t}$ to be the set of all n for which $t_k(n) < t$ and $V_{k,t}$ to be the set of all n for which $t_k(n) = t$. The result obtained for $U_{k,t}$ is given in Theorem 1 and that for $V_{k,t}$ in Theorem 2. It will be noted that $U_{k,k}$ is the set of all positive integers while $U_{k,1}$ is the set of k -th powers. More generally, $U_{k,t}$ is the set of all n whose k -segment is a t -free integer (an integer whose largest t -th power divisor is 1). Note moreover that the sets $V_{k,t}$ contain no k -th power integers.

We mention here a single special case of Theorem 1. Let $H_k(x)$ denote the number of $n \leq x$ contained in $H_k = U_{k,2}$, that is, $H_k(x)$ is the number of those n , not exceeding x , whose k -segment is square-free. Then for $k > 2$,

$$H_k(x) = \left(\frac{\zeta(k)}{\zeta(2)} \right) x + O(\sqrt{x}), \quad (1.2)$$

*) This research was supported in part by the National Science Foundation grant G-19012.