

Ein Verfahren zum Minimieren einer Funktion bei eingeschränktem Variationsbereich der Parameter*

G. ZETTL

Eingegangen am 12. Dezember 1968

Summary. An algorithm is proposed for minimization of a function Φ of many variables, which is capable of handling constraints in the form of linear inequalities among the variables and which has second order convergence properties. It is assumed that Φ and its partial derivatives can be computed. The method is related to that suggested by Fletcher and Powell. However, a different algorithm is used to approximate the inverse of the Hessian matrix which enables the insertion of constraints. Examples show that convergence is as good as in the method mentioned above.

Zusammenfassung. Es wird eine neue, quadratisch konvergierende Methode vorgeschlagen zur Minimierung einer Funktion Φ von mehreren Veränderlichen, bei der den Parametern Nebenbedingungen in Form linearer Ungleichungen auferlegt werden können. Neben dem Funktionswert Φ werden die partiellen Ableitungen von Φ nach den Parametern benötigt. Die Ableitungen zweiter Ordnung werden indirekt vom Verfahren berücksichtigt. Die Methode ist mit der von Fletcher und Powell angegebenen verwandt. Zum Aufbau der Inversen der Matrix der zweiten Ableitungen wird jedoch ein anderer, neuer Algorithmus benutzt, der das Einführen von Nebenbedingungen ermöglicht. Beispiele zeigen, daß die neue Methode in ihrem Konvergenzverhalten der erwähnten ebenbürtig ist.

1. Einführung

Optimierungsprobleme der Technik lassen sich oft als eine Minimierungsaufgabe formulieren. Die Optimallösung deckt sich dabei mit dem Minimum einer Fehlerquadratsumme. Frei laufende Parameter können zu unerwünschten, weil nicht realisierbaren, Lösungen führen. Ziel dieser Arbeit ist es, eine Methode vorzustellen, die folgende Aufgabe löst:

Man bestimme das Minimum der Funktion

$$\Phi = \Phi(x_1, x_2, \dots, x_n), \quad (1)$$

wobei die Parameter x_1, x_2, \dots, x_n den linearen Ungleichungen

$$b_{k0} + \sum_{i=1}^n b_{ki} x_i \geq 0, \quad k=1, 2, \dots, l; \quad l \geq n \quad (2)$$

genügen sollen.

* Mitteilung aus dem Zentrallaboratorium für Nachrichtentechnik der Siemens AG, München.

Ein Lösungsweg, mit dem sich diese Aufgabe bewältigen läßt, wurde von Fiacco und McCormick in [3] und [4] vorgeschlagen und von Lasdon und Waren in [6] angewandt. Dieser Lösungsweg gestattet die Einführung allgemeiner Nebenbedingungen, beschränkt sich also nicht auf Nebenbedingungen bezüglich der Parameter in der Form (2), erfordert dafür aber einen vergleichsweise großen Rechenaufwand.

Das hier vorzustellende Verfahren ist mit dem von Fletcher und Powell [5] angegebenen verwandt. Es arbeitet auch mit einer Näherung zweiter Ordnung für die Funktion und benötigt dazu Funktions- und Gradientenwerte.

Die Glieder zweiter Ordnung werden durch einen anderen Algorithmus als in [5] berücksichtigt. Dieser neue Algorithmus stellt keine Orthogonalitätsforderungen an den Iterationsablauf und gestattet deshalb das Einführen von Nebenbedingungen.

Konvergenzvergleiche mit [5] lassen sich nur bei frei laufenden Parametern durchführen. In den gerechneten Testbeispielen konvergierte die hier zu beschreibende Methode etwa ebenso gut wie die von Fletcher und Powell.

2. Verwendete Symbole

In dieser Arbeit wird für Vektoren und Matrizen die Diracsche Schreibweise [2] verwendet. Elemente dieser Schreibweise sind der senkrechte Strich und die nach links bzw. rechts geöffnete Ecke. Dabei bedeuten:

$\langle x \rangle$	Skalar
$ x\rangle$	Spaltenvektor
$\langle x $	Zeilenvektor
$ x $	Matrix.

Die Art einer zusammengesetzten Größe läßt sich einfach aus den links und rechts von ihr stehenden Strichen oder Ecken entnehmen, z.B.

$\langle x y\rangle$	Skalarprodukt
$ x\rangle\langle y $	dyadisches Produkt (Matrix).

Die im folgenden verwendeten Größen seien hier vorweg zusammengestellt:

α	Parameter der eindimensionalen Minimumsuche
b_{ij}	Koeffizient der Nebenbedingungen
$ B $	Rechteckige Matrix mit Elementen b_{ij}
Φ	Zu minimierende Funktion
F	Nach den Gliedern zweiter Ordnung abgebrochene Taylorentwicklung von Φ
$ g(x)\rangle$	Gradientenvektor an der Stelle $ x\rangle$
$ \Delta g\rangle$	Differenz zweier Gradientenvektoren
$ H_i $	Iterationsmatrix, Näherung für $ K^{-1} $

$ K $	Matrix der zweiten Ableitungen
$ A\rangle$	Vektor der Lagrange-Parameter
m	Anzahl der erfaßten Randebenen
n	Anzahl der Parameter
Q	Unter Nebenbedingungen zu minimierende Funktion
$ r\rangle$	Suchstrahl
$ s\rangle$	Schrittorschlag ohne Nebenbedingungen
$ u\rangle$	Schrittorschlag unter Nebenbedingungen
$ x\rangle$	Ortsvektor
$ \Delta x\rangle$	Differenz zweier Ortsvektoren.

3. Behandlung der Nebenbedingungen

Das Verfahren beginnt mit einer im n -dimensionalen Variationsbereich liegenden Näherung. Variationsraum ist ein von ebenen Flächen berandeter Körper, in dem sich der Parametervektor frei bewegen darf. Zunächst zielen die Parameteränderungen ohne einschränkende Nebenbedingungen auf eine möglichst starke Reduktion der Funktion Φ . Dabei kann es geschehen, daß Berandungsebenen durchstoßen werden. Das Minimum von Φ liegt dann offenbar außerhalb des Körpers. Im Körper nimmt Φ seinen kleinsten Wert auf dem Rand an, und es gilt, dieses Minimum auf dem Rand zu bestimmen. Die Reduktion erfolgt jetzt unter den durch die Berandung gegebenen Nebenbedingungen. Dabei ist normalerweise nur ein Teil der Begrenzungsebenen im Spiel. Die restlichen Randebenen sind in den vorausgehenden Iterationsschritten nicht angesprochen worden und bleiben deshalb unbeachtet. Im weiteren Verlauf der Iteration kann sich herausstellen, daß das wirkliche Minimum der Funktion doch im Innern des Körpers liegt, nur der Weg zu diesem Minimum führte über den Rand. Dann sind die betreffenden Randnebenbedingungen wieder freizugeben.

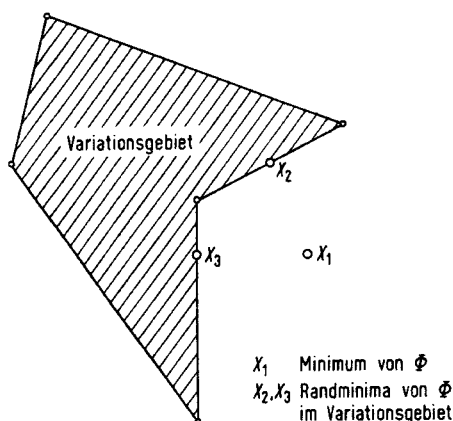


Fig. 1. Nichtkonvexes Variationsgebiet mit zwei Randminima

Daß die Nebenbedingungen (2) keinen Widerspruch enthalten dürfen, versteht sich von selbst. Wenn sie einen nicht konvexen Körper beschreiben (konvex: je zwei Punkte im Innern des Körpers müssen sich durch eine ganz im Innern des Körpers liegende Strecke verbinden lassen), dann ist zu beachten, daß auch bei einer im Gesamttraum konvexen Funktion Φ das Minimum nicht mehr eindeutig zu sein braucht. In Fig. 1 habe Φ ein Minimum außerhalb des Variationsgebietes. Im Variationsbereich existieren zwei verschiedene Minima.

4. Die Methode

Kern des Verfahrens ist eine nach den Gliedern zweiter Ordnung abgebrochene Taylorreihe, die die zu minimierende Funktion annähert:

$$F = f_0 + \langle g(x_0) | x - x_0 \rangle + \frac{1}{2} \langle x - x_0 | K | x - x_0 \rangle. \quad (3)$$

Diese Näherungsfunktion wird iterativ aufgebaut. Die Taylorentwicklung einer Funktion hängt ab von dem Punkt, um den entwickelt wird. Bei der Iteration bewegt sich der Parametervektor, ausgehend von der Näherung, durch den Raum zum Minimum. Die Taylorentwicklung der Funktion wandelt sich dabei von der Entwicklung um die Näherung zur Entwicklung um die Lösung. Bestimmendes Element der Näherungsfunktion ist die Matrix $|K|$ oder deren Inverse $|K^{-1}|$. Für diese Matrix wird eine Näherung ermittelt aus einander zugeordneten Änderungen des Parameter- und Gradientenvektors. Jeder Iterationsschritt liefert solch ein Paar von Vektoren. Für eine quadratische Funktion sind zur Bestimmung der Matrix n linear unabhängige Vektorpaare erforderlich. Im allgemeinen bestimmen die letzten n Vektorpaare die quadratische Näherung. In Gebieten zäher Konvergenz und in der Umgebung des Minimums wird die Matrix automatisch genauer aufgebaut als in anderen Gebieten. Dies geschieht dadurch, daß der Parametervektor mehrere Schritte lang in diesem Gebiet verweilt und mehrere Vektorpaare aus diesem Gebiet zum Aufbau der Matrix benutzt werden.

Nebenbedingungen beeinträchtigen die Konvergenz des Verfahrens nicht. Der Verbesserungsvektor unter Nebenbedingungen ergibt sich aus dem bei frei laufenden Parametern durch ein Korrekturglied.

Der Übergang von frei laufenden Parametern zur Iteration unter Nebenbedingungen geschieht ohne Informationsverlust, denn die erreichte Näherungsfunktion wird beibehalten und weiterbenutzt.

Die Näherungsmatrix strebt gegen die Inverse der Matrix der zweiten Ableitungen und braucht nicht unbedingt positiv definit zu sein wie in der Methode von Fletcher und Powell [5]. Es ergibt sich eine echte Näherung für die Funktion in der Umgebung des momentanen Parametervektors, auch wenn ein Sattelpunkt vorhanden ist. Mit Hilfe dieser Matrix wird näherungsweise der Punkt bestimmt, in dem der Gradient verschwindet. Dies ist nicht nur im Minimum, sondern auch im Sattelpunkt der Fall. In der Nähe eines Minimums zielen die Schrittvorschläge zu diesem hin. Bei einem Sattelpunkt könnte ein derartiger Schrittvorschlag in Richtung wachsender Funktionswerte zeigen. Er würde zum Sattelpunkt statt zum Minimum führen. In diesem Fall wird, damit man vom Sattel wegkommt, ein Schritt vorgeschlagen in Richtung des Eigenvektors, der dem kleinsten Eigenwert der Näherungsmatrix für die zweiten Ableitungen zugeordnet ist.

Zum Aufbau der Iterationsmatrix wird ein anderer Algorithmus verwendet als in [5]. Dieser Algorithmus verlangt nicht die Beachtung von Orthogonalitätsforderungen. Die Schritte können beliebig sein; für den Aufbau der Matrix ist lediglich erforderlich, daß die Vektorpaare den ganzen Parameterraum erfassen, also linear unabhängig sind.

Als Startmatrix kann die Einheitsmatrix verwendet werden. Bei guten Näherungen empfiehlt es sich, aus n Prohebewegungen, z.B. längs den Koordinatenachsen, die Matrix vorher aufzubauen.

Damit das Verfahren stabil ist, wird bei jedem Schritt eine eindimensionale Minimumsuche vorgenommen. Das Minimum braucht hierbei nicht genau bestimmt zu werden.

5. Der Algorithmus

Zur Funktion

$$F(x) = f_0 + \langle g(x_0) | x - x_0 \rangle + \frac{1}{2} \langle x - x_0 | K | x - x_0 \rangle \quad (3)$$

gehört der Gradient

$$|g(x)\rangle = |g(x_0)\rangle + |K| x - x_0\rangle. \quad (4)$$

Im Minimum der Funktion verschwindet der Gradient. Für den Zuschlagvektor $|s\rangle$, der zum Minimum führt, folgt:

$$|s\rangle = |x - x_0\rangle = -|K^{-1}|g(x_0)\rangle. \quad (5)$$

Haben einer oder mehrere der Parameter den Rand erreicht, so ist das Minimum von F unter der Nebenbedingung

$$|b_0\rangle + |B|x\rangle = 0 \quad (6)$$

zu bestimmen. Gl. (6) geht aus dem Ungleichungssystem (2) hervor. Sind m Randbedingungen zu beachten, dann ist $|b_0\rangle$ ein m -dimensionaler Vektor und $|B|$ eine rechteckige Matrix aus n Spalten und m Zeilen.

Für die Minimumsuche unter Nebenbedingungen bietet sich die Lagrange-Methode an. Hierbei ist das Minimum der Funktion

$$\begin{aligned} Q(x, \lambda) &= F(x) + \langle \lambda [|b_0\rangle + |B|x\rangle] \\ &= f_0 + \langle g(x_0) | x - x_0 \rangle + \frac{1}{2} \langle x - x_0 | K | x - x_0 \rangle + \langle \lambda [|b_0\rangle + |B|x\rangle] \end{aligned} \quad (7)$$

zu finden. $|\lambda\rangle$ ist dabei der m -dimensionale Vektor der Lagrange-Parameter.

Für $|\lambda\rangle$ und $|x\rangle$ gelten neben der Gl. (6) die Beziehungen

$$|b_0\rangle + |B|x_0\rangle = 0, \quad (8)$$

$$|g(x_0)\rangle + |K|x - x_0\rangle + |B^t|\lambda\rangle = 0. \quad (9)$$

Aus diesen Gleichungen läßt sich der zum Minimum führende Zuschlag $|u\rangle$ entnehmen:

$$\text{Aus (6) und (8) folgt } |B|x - x_0\rangle = 0. \quad (10)$$

$$\text{Aus (9) } |x - x_0\rangle = -|K^{-1} [|g(x_0)\rangle + |B^t|\lambda\rangle]. \quad (11)$$

$$(11) \text{ in } (10): \quad |B|K^{-1}|g(x_0)\rangle + |B|K^{-1}|B^t|A\rangle = 0.$$

$$\text{Daraus} \quad |A\rangle = - [|B|K^{-1}|B^t|]^{-1} |B|K^{-1}|g(x_0)\rangle. \quad (12)$$

$$(12) \text{ in } (11):$$

$$|x - x_0\rangle = |u\rangle = - |K^{-1}|g(x_0)\rangle + |K^{-1}|B^t [|B|K^{-1}|B^t|]^{-1} |B|K^{-1}|g(x_0)\rangle$$

oder mit (5)

$$|u\rangle = |s\rangle + |K^{-1}|B^t [|B|K^{-1}|B^t|]^{-1} |B|K^{-1}|g(x_0)\rangle. \quad (13)$$

Wie man sieht, besteht der Zuschlag $|u\rangle$ unter Nebenbedingungen aus dem Zuschlag $|s\rangle$ bei freier Iteration und einem Korrekturglied. Gl. (13) verlangt keine vollständige Matrixinversion, wie es zunächst scheint. Es genügt, wenn das Gleichungssystem

$$|B|K^{-1}|B^t|z\rangle = |B|K^{-1}|g(x_0)\rangle \quad (14)$$

nach $|z\rangle$ aufgelöst wird.

Die Arbeitsweise des Verfahrens sei an Hand des Flußdiagramms Fig. 2 erläutert.

Bei guten Ausgangsnäherungen ist es vorteilhaft, wenn vor Iterationsbeginn die Matrix $|K^{-1}|$, die Inverse der Matrix der zweiten Ableitungen von F und damit auch von Φ , bestimmt wird. Auf diese Weise lassen sich Schritte und Funktionswertberechnungen einsparen. Zum Aufbau der Matrix wird der Gradient der Funktion Φ in $n+1$ Punkten der engeren Umgebung der Ausgangsnäherung benötigt. Die Matrix gewinnt man wie folgt:

Man bildet

$$|\Delta x_i\rangle = |x_i\rangle - |x_0\rangle, \quad (15)$$

$$|\Delta g_i\rangle = |g_i\rangle - |g_0\rangle, \quad (16)$$

$$|y_i\rangle = |\Delta x_i\rangle - |H_{i-1}|\Delta g_i\rangle, \quad (17)$$

$$|H_i| = |H_{i-1}| + \frac{|y_i\rangle\langle y_i|}{\langle y_i|\Delta g_i\rangle}. \quad (18)$$

Der Index i läuft dabei von 1 bis n . Als $|\Delta x_i\rangle$ können z.B. Änderungen längs der Koordinatenachsen gewählt werden. $|H_0|$ kann eine beliebige symmetrische Matrix sein, etwa die Einheitsmatrix.

$|H_n|$ stimmt mit $|K^{-1}|$ der Taylorentwicklung überein. Näheres über diese Vorschrift für den Matrixaufbau findet sich in Abschnitt 6.

Ist die Ausgangsnäherung schlecht, dann bringt ein Matrixaufbau am Anfang keine Vorteile. Er kostet nur unnötig Zeit. In diesem Fall startet man am besten mit der Einheitsmatrix.

Die vorliegende Matrix $|H|$ braucht nicht unbedingt positiv definit zu sein, und es kann daher geschehen, daß der Schrittvorschlag $|s\rangle$ (oder $|u\rangle$) in Richtung wachsender Funktionswerte zeigt. Die Matrix enthält dann negative Eigenwerte, ein Zeichen dafür, daß ein Sattelpunkt in der Nähe liegt und die Näherung nicht für das gesuchte Minimum gilt. In diesem Fall ist es am günstigsten, wenn der Schrittvorschlag in das vom Sattelpunkt am steilsten abwärts führende Tal weist. Dieses Tal läuft vom Sattel aus in Richtung des zum kleinsten Eigenwert ge-

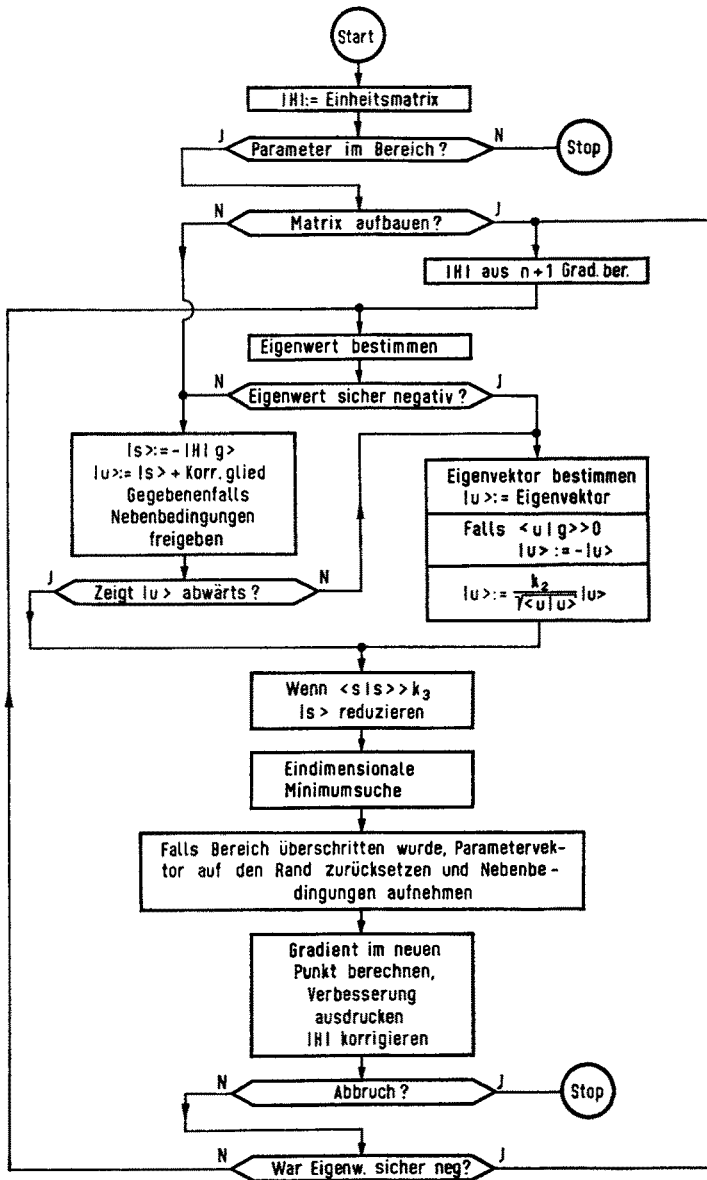


Fig. 2. Flußdiagramm

hörenden Eigenvektors der Matrix $|K|$, sofern $|K|$ die Umgebung dieses Sattels richtig beschreibt. Für $|K^{-1}|$ ist dies der größte negative Eigenwert.

Unter Nebenbedingungen tritt an Stelle von $|K^{-1}|$ die Matrix

$$|C| = |K^{-1}| - |K^{-1}| B' [B | K^{-1}| B']^{-1} B |K^{-1}|, \quad (19)$$

wie man aus (13) leicht sieht.

Diese Matrix ist so aufgebaut, daß sie Änderungen des Gradienten in solche Änderungen des Parametervektors transformiert, die den Rand des Bereichs, also die m erfaßten Begrenzungsflächen, nicht verlassen. Alle Vektoren, die zu diesen Randebenen orthogonal sind, werden von $|C|$ in Null transformiert. $|C|$ ist daher singular mit dem Rangabfall m und hat den m -fachen Eigenwert 0.

Sind die übrigen $n - m$ Eigenwerte sämtlich positiv, dann wird nach (5) $|s\rangle$ errechnet. An die Stelle von $|K^{-1}|$ tritt dabei in dieser Gleichung natürlich $|H|$. $|H|$ stimmt nur bei vorausgegangenem Matrixaufbau oder gegen Ende der Iteration mit $|K^{-1}|$ überein. Unter Nebenbedingungen ist an $|s\rangle$ noch eine Korrektur vorzunehmen. Bevor diese angebracht wird, ist zu untersuchen, ob $|s\rangle$ von einer oder mehreren der Berandungsebenen zurück in den Körper zeigt. Ist dies der Fall, dann werden die betreffenden Nebenbedingungen probeweise freigegeben. Weist auch der korrigierte Vektor $|u\rangle$ nicht aus dem Körper heraus, dann war die Freigabe berechtigt, ansonsten war sie verfrüht, und der Vektor muß wieder rückgestellt werden.

Die Eigenwerte von $|H|$ und $|C|$ können nur mit beschränkter Genauigkeit ermittelt werden. Es ist daher sinnvoll, auch bei negativen Eigenwerten, die an der Grenze des positiven Bereiches liegen, die Matrix $|H|$ oder $|C|$ zunächst einmal als positiv anzunehmen. Die Matrix wird als nicht positiv definit angesehen, wenn sie Eigenwerte in dem Intervall $k_0 < \lambda < k_1$ besitzt. Die Größen k_0 und k_1 sind dabei entsprechend der zu erwartenden Ungenauigkeit zu wählen (z.B. $k_0 = -10^{-2}$, $k_1 = 1/k_0$). Zeigt der Schrittvorschlag $|u\rangle$ (bzw. $|s\rangle$, wenn keine Nebenbedingungen vorliegen) abwärts, ist also $\langle u | -g \rangle > 0$, dann ist er brauchbar, ansonsten wird wie bei negativem Eigenwert vorgegangen.

Bei negativen Eigenwerten, genauer, wenn ein Eigenwert in das Intervall $k_0 < \lambda < k_1$ fällt, wird die Richtung des zum kleinsten dieser Eigenwerte zugehörigen Eigenvektors bestimmt. Damit dieser Vektor in Richtung fallender Funktionswerte weist, muß gegebenenfalls noch sein Vorzeichen umgekehrt werden. Der Eigenvektor gibt nur eine Richtung, sein Betrag liegt nicht fest. In dieser Richtung wird ein kleiner Schritt vom Betrag k_2 vorgeschlagen (z.B. $k_2 = 10^{-2}$).

Komplizierte Funktionen besitzen mehrere Minima. Übersteigt der Betrag von $|s\rangle$ (bzw. $|u\rangle$) eine Schranke k_3 , dann wird er auf diese Schranke reduziert. Auf diese Weise soll erreicht werden, daß die Iteration in jenes Minimum mündet, das der Rohnäherung am nächsten liegt. Man verhindert so, daß die Iteration in ein „Nachbartal“ springt, das zu einem anderen Minimum führt.

Der Schrittvorschlag liegt damit fest und zeigt abwärts. Um Stabilität und Konvergenz des Verfahrens zu sichern, wird längs des durch den Schrittvorschlag gegebenen Strahls das Minimum bestimmt. Dieses Minimum braucht nicht genau ermittelt zu werden, denn es bestehen keine Orthogonalitätsforderungen wie in [5]. In Abschnitt 7 wird eine Methode zur Bestimmung dieses Minimums näher beschrieben.

Im Anschluß an diese eindimensionale Minimierung wird untersucht, ob der Parametervektor noch im vorgeschriebenen Körper liegt. Wenn nicht, wird er auf die Berandungsfläche zurückgesetzt und die durchstoßene Randebene als Nebenbedingung aufgenommen.

Der Schritt ist jetzt ausgeführt. Im neuen Punkt wird der Gradient berechnet und aus der Differenz des neuen und des alten Gradienten und der Parameter-

änderung die Matrix $|H|$ korrigiert:

$$|\Delta x\rangle = |x_{j+1}\rangle - |x_j\rangle; \quad (15a)$$

$$|\Delta g\rangle = |g_{j+1}\rangle - |g_j\rangle; \quad (16a)$$

$$|y\rangle = |\Delta x\rangle - |H_j|\Delta g\rangle; \quad (17a)$$

$$|H_{j+1}| = |H_j| + \frac{|y\rangle\langle y|}{\langle y|\Delta g\rangle}. \quad (18a)$$

Für die Korrektur findet der bereits zum Matrixaufbau benutzte und in Abschnitt 6 näher beschriebene Algorithmus Verwendung.

Im nächsten Programmabschnitt wird geprüft, ob die Kriterien für den Abbruch der Iteration erfüllt sind. Als Abbruchkriterien dienen dabei die größte Parameteränderung und der Betrag der Gradientenkomponente senkrecht zu den erfaßten Randebenen.

Ist das Minimum noch nicht erreicht, dann wird der Zyklus erneut durchlaufen. War die Näherungsfunktion eindeutig eine Sattelfläche (λ im Intervall $k_0 < \lambda < k_1$) und wurde der Schritt parallel zum Tal dieser Fläche ausgeführt, dann beginnt der neue Zyklus mit einem Matrixaufbau. Diesem Vorgehen liegen folgende Überlegungen zugrunde: Bei einer quadratischen Funktion würde die Funktion längs dieser Richtung beständig abnehmen. Der Schritt in Eigenvektorrichtung führt aber bei der eindimensionalen Minimumsuche bis zu einem Minimum und damit aus der Umgebung heraus, in der die Näherungsfunktion F die zu minimierende Funktion Φ beschrieb. In der neuen Umgebung gilt die Näherung nicht mehr. Um in dieser Umgebung die Iteration mit einer brauchbaren Näherungsfunktion fortführen zu können, wird die Matrix für diese Umgebung neu bestimmt. Ein Matrixaufbau nach den Gln. (15)–(18) kostet weniger Zeit als n Schritte, in denen neben den Gradientenberechnungen noch zusätzlich Funktionswerte für die eindimensionalen Minimierungen berechnet werden müssen.

6. Beweis der Vorschrift für den Matrixaufbau

Es seien n Vektorpaare gegeben,

$$|\Delta g_i\rangle, \quad |\Delta x_i\rangle; \quad i = 1, n,$$

und es existiere eine symmetrische Matrix $|K^{-1}|$ mit der Eigenschaft

$$|K^{-1}|\Delta g_i\rangle = |\Delta x_i\rangle; \quad i = 1, n. \quad (20)$$

Diese Matrix ist aus einer beliebigen symmetrischen Rohmatrix $|H_0|$ schrittweise aufzubauen:

$$|H_j| = |H_{j-1}| + |\Delta H_j|. \quad (21)$$

Die Matrix $|H_j|$ soll die ersten j Vektoren $|\Delta g_i\rangle$ ($i = 1, j$) in die entsprechenden $|\Delta x_i\rangle$ transformieren. Sofern die $|\Delta g_i\rangle$ linear unabhängig sind, ist dann nach n Schritten

$$|H_n| \equiv |K^{-1}|. \quad (22)$$

Der sukzessive Aufbau der Matrix $|K^{-1}|$ ist im Iterationsverfahren deshalb vorteilhaft, weil sich die Matrix der Ableitungen zweiter Ordnung, deren Inverse ja

$|K^{-1}|$ ist, bei beliebigen Funktionen während der Iteration ändert. Für eine Form zweiter Ordnung ist diese Matrix vom Ort unabhängig und konstant.

Gl. (21) ist bereits der Lösungsansatz für die gestellte Aufgabe. Ehe auf die von $|H_j|$ zu fordernden Eigenschaften eingegangen wird, zunächst noch einige allgemeine Betrachtungen.

Durch Linearkombinationen von $j-1$ linear unabhängigen Vektoren $|\Delta g_i\rangle$ wird ein $(j-1)$ -dimensionaler Unterraum $R_{g,j-1}$ des n -dimensionalen Vektorraumes R aufgespannt. $|K^{-1}|$ bildet diesen Raum auf einen Raum $R_{x,j-1}$ ab. In $R_{x,j-1}$ liegen die Vektoren $|\Delta x_i\rangle$, $i=1, j-1$.

Nimmt man an, $|H_{j-1}|$ transformiere die Vektoren $|\Delta g_i\rangle$, $i=1, \dots, j-1$ in der vorgeschriebenen Weise, dann transformiert $|H_{j-1}|$ den Raum $R_{g,j-1}$ in den Raum $R_{x,j-1}$. $|\Delta H_j|$ muß $R_{g,j-1}$ in Null transformieren, denn er wird ja von $|H_{j-1}|$ bereits richtig transformiert. Für die Verbesserungsmatrix setzt man am einfachsten eine Matrix vom Range Eins an. Der Eigenvektor dieser Matrix mit nicht verschwindendem Eigenwert sei $|y_j\rangle$. $|y_j\rangle$ muß orthogonal sein zu $R_{g,j-1}$. $R_{g,j-1}$ schneidet aus dem Gesamtraum R einen Teilraum heraus. Der verbleibende Restraum wird von $|H_{j-1}|$ in beliebiger Weise transformiert. $|\Delta g_j\rangle$ hat eine Komponente in diesem Restraum, wenn es von den $|\Delta g_i\rangle$, $i=1, \dots, j-1$ nicht linear abhängt. Diese Komponente transformiert $|H_{j-1}|$ normalerweise falsch. $|\Delta H_j|$ soll diesen Fehler ausgleichen.

Die Differenzmatrix

$$|D_j| = |K^{-1}| - |H_{j-1}| \quad (23)$$

bildet $R_{g,j-1}$ in Null ab.

$|D_j|$ ist symmetrisch und hat deshalb orthogonale Eigenvektoren. $|D_j|$ ist auch singulär mit dem Rangabfall $j-1$. Die Eigenvektoren von $|D_j|$ mit verschwindendem Eigenwert spannen gerade den Raum $R_{g,j-1}$ auf. Denkt man sich $|D_j|$ nach den Eigenvektoren $|e_k\rangle$ zerlegt,

$$|D_j| = \sum_{k=1}^n \frac{|e_k\rangle\langle e_k|}{\langle e_k|e_k\rangle} \lambda_k, \quad (24)$$

und multipliziert man diese Zerlegung mit einem beliebigen Vektor, dann ist unmittelbar einzusehen, daß der Produktvektor keine Komponente in $R_{g,j-1}$ haben kann, weil die Eigenvektoren orthogonal sind und jene Eigenvektoren, die in $R_{g,j-1}$ transformieren, den Eigenwert Null haben. Das Produkt

$$|D_j|\Delta g_j\rangle = |\Delta x_j\rangle - |H_{j-1}|\Delta g_j\rangle, \quad (25)$$

das gerade den Fehler liefert, um den $|H_{j-1}|$ den Vektor $|\Delta g_j\rangle$ falsch transformiert, ist folglich orthogonal zu $R_{g,j-1}$. Es hat damit alle Eigenschaften, die vom Eigenvektor $|y_j\rangle$ von $|\Delta H_j|$ zu fordern sind:

$$|y_j\rangle = |\Delta x_j\rangle - |H_{j-1}|\Delta g_j\rangle. \quad (26)$$

Die Korrekturmatrix $|\Delta H_j|$ ist also bis auf einen Faktor γ bestimmt:

$$|\Delta H_j| = \gamma |y_j\rangle\langle y_j|. \quad (27)$$

Aus der Forderung

$$|H_j|\Delta g_j\rangle = |\Delta x_j\rangle$$

ergibt sich mit (21), (26) und (27)

$$\gamma = \frac{1}{\langle y_j | \Delta g_j \rangle} \quad (28)$$

und somit

$$|\Delta H_j| = \frac{|y_j\rangle \langle y_j|}{\langle y_j | \Delta g_j \rangle}. \quad (29)$$

Die Rohmatrix $|H_0|$ wird normalerweise keinen der Vektoren $|\Delta g_i\rangle$ richtig transformieren. $|H_1|$ transformiert aber dann $|\Delta g_1\rangle$ wie vorgeschrieben in $|\Delta x_1\rangle$. Durch Schluß von n auf $n+1$ ergibt sich: $|H_i|$ transformiert die ersten i Vektoren richtig, und $|H_n|$ stimmt schließlich mit $|K^{-1}|$ überein, denn es transformiert den ganzen Raum R .

$|H_j|$ braucht nicht positiv definit zu sein. $|H_j|$ strebt gegen $|K^{-1}|$. Besitzt $|K^{-1}|$ negative Eigenwerte, dann hat auch $|H_n|$ welche.

Ein kleines Beispiel soll die Arbeitsweise des Algorithmus zeigen. Gegeben sei die Matrix

$$|K| = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 3 & 8 & 14 \end{vmatrix};$$

zwischen $|\Delta g\rangle$ und $|\Delta x\rangle$ gelte die Beziehung

$$|\Delta g\rangle = |K|\Delta x\rangle.$$

Anfangsmatrix sei die Einheitsmatrix. Die Tabelle zeigt den Rechengang bei willkürlich gewählten $|\Delta x_i\rangle$. Nach drei Schritten erhält man die Inverse von $|K|$.

Tabelle. Zahlenbeispiel für Matrixaufbau

i	H_{i-1}	$ \Delta x_i\rangle$	$ \Delta g_i\rangle$	$ y_i\rangle$
1	$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$	$\begin{vmatrix} 1 \\ 0 \\ 0 \end{vmatrix}$	$\begin{vmatrix} 1 \\ 2 \\ 3 \end{vmatrix}$	$\begin{vmatrix} 0 \\ -2 \\ -3 \end{vmatrix}$
2	$\frac{1}{13} \begin{vmatrix} 13 & 0 & 0 \\ 0 & 9 & -6 \\ 0 & -6 & 4 \end{vmatrix}$	$\begin{vmatrix} 0 \\ 1 \\ 0 \end{vmatrix}$	$\begin{vmatrix} 2 \\ 5 \\ 8 \end{vmatrix}$	$\frac{1}{13} \begin{vmatrix} -26 \\ 16 \\ -2 \end{vmatrix}$
3	$\frac{1}{3} \begin{vmatrix} 16 & -8 & -1 \\ -8 & 7 & -2 \\ 1 & -2 & 1 \end{vmatrix}$	$\begin{vmatrix} 0 \\ 0 \\ 1 \end{vmatrix}$	$\begin{vmatrix} 3 \\ 8 \\ 14 \end{vmatrix}$	$\frac{2}{3} \begin{vmatrix} 1 \\ 2 \\ 1 \end{vmatrix}$
4	$\begin{vmatrix} 6 & -4 & 1 \\ -4 & 5 & -2 \\ 1 & -2 & 1 \end{vmatrix}$		—	

7. Minimumsuche längs des Strahls

Das Verfahren, das zur Bestimmung des Minimums längs des Strahls

$$|r\rangle = \alpha |s\rangle \quad (30)$$

verwendet wurde, setzt sich aus drei Teilen zusammen, nämlich:

a) einem Intervallaufweitungsverfahren, bei dem durch Intervallvergrößerung längs des betrachteten Strahls ein Intervall bestimmt wird, in dem das Minimum liegt;

b) einem Intervallschachtelungsverfahren, das dieses Intervall auf einen vorgegebenen Wert einengt;

c) einem Interpolationsverfahren, das in dem durch Schachtelung gewonnenen Intervall das Minimum durch Interpolation bestimmt.

Die einzelnen Teile der Minimumsuche sollen zusammenpassen. Der komplizierteste Teil ist das Schachtelungsverfahren. Es legt die übrigen Teile weitgehend fest und soll deshalb zuerst besprochen werden.

Ist in einem Intervall die Ordinate eines inneren Punktes kleiner als die der Randpunkte, dann existiert ein Minimum in dem Intervall. (Die Funktion ist dabei als konvex vorauszusetzen.) Kennt man in dem Intervall noch einen weiteren Punkt, dann läßt sich ein Teilintervall abspalten, in dem das gesuchte Minimum enthalten ist.

In Fig. 3 liegt das Minimum im Intervall (a_1, a_4) . Die beiden Intervalle (a_1, a_3) , (a_2, a_4) überdecken das Intervall (a_1, a_4) .

Das Minimum der Funktion muß also in mindestens einem der beiden Teilintervalle liegen. In welchem es liegt, hängt davon ab, welcher der beiden mittleren Punkte die kleinere Ordinate hat.

Aus dem Intervall (a_1, a_4) entsteht so das Intervall (a_1, a_3) oder (a_2, a_4) . Berechnet man in dem Teilintervall einen weiteren Punkt, dann kann aus diesem ein noch kleineres Intervall herausgespalten werden.

Die Punkte a_2, a_3 sollten das Intervall (a_1, a_4) gleichmäßig aufteilen, damit Ungenauigkeiten in der Funktionsberechnung möglichst wenig stören und die Intervalllängen schnell abnehmen. Fortgesetztes Schachteln sollte die Teilverhält-

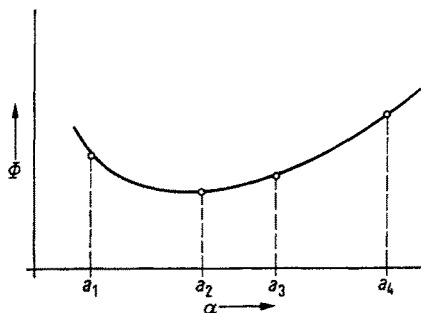


Fig. 3. Intervallschachtelung bei der eindimensionalen Minimumsuche

nisse erhalten. Aus diesen Überlegungen folgt der Ansatz:

$$\frac{a_3 - a_1}{a_4 - a_1} = \frac{a_2 - a_1}{a_3 - a_1}, \quad (31)$$

$$(a_2 - a_1) + (a_4 - a_2) = (a_4 - a_1), \quad (32)$$

$$(a_4 - a_2) = (a_3 - a_1). \quad (33)$$

Die Gln. (31), (32) und (33) führen zu der Lösung

$$(a_3 - a_1) = \left[-\frac{1}{2} \pm \frac{1}{2}\sqrt{5}\right] (a_4 - a_1). \quad (34)$$

$(a_3 - a_1)$ soll positiv sein. Somit bleibt nur die Lösung:

$$(a_3 - a_1) = \frac{\sqrt{5} - 1}{2} (a_4 - a_1). \quad (35)$$

Für die andere Teilstrecke folgt damit aus (31)

$$(a_2 - a_1) = \frac{3 - \sqrt{5}}{2} (a_4 - a_1). \quad (36)$$

Die Gln. (35) und (36) besagen, daß die Strecke $\overline{a_1 a_3}$ durch die Punkte a_2 und a_4 innen und außen harmonisch geteilt wird. Im angegebenen Beispiel läßt sich aus dem Intervall (a_1, a_4) das Teilintervall (a_1, a_3) herauspalten. Wurden a_2 und a_3 in obiger Weise bestimmt, dann läßt sich (a_1, a_3) durch einen weiteren Punkt zu einem Vier-Punkte-Intervall ergänzen, das wieder harmonisch aufgeteilt ist. Nach vier Schritten, also nach Berechnung von vier weiteren Funktionswerten, ist das ursprüngliche Intervall auf etwa ein Zehntel seiner Breite zusammengeschrumpft. Die Schachtelung wird abgebrochen, wenn die Intervallbreite einen vorgegebenen Wert unterschreitet. Dieses Schachtelverfahren ist eine einfache und gute Näherung für die Fibonacci-Schachtelung [1].

Dem Schachtelverfahren ist das Aufweitungsverfahren vorgeschaltet. Bei der Aufweitung soll sich ein harmonisch geteiltes Intervall ergeben, an das die Schachtelung unmittelbar angeschlossen werden kann. Für die Aufweitung wurde deshalb folgende Methode gewählt:

Man beginnt mit zwei Punkten a_0 und a_1 und ergänzt diese durch einen dritten Punkt a_2 , so daß das Intervall (a_0, a_2) von a_1 harmonisch geteilt wird. Im erstellten Programm haben die Punkte a_0 den Wert $\alpha = 0$ und a_2 den Wert $\alpha = 1$.

Aus drei Punkten läßt sich entscheiden, ob im vorliegenden Intervall ein Minimum enthalten ist. Ist $f(a_2) < f(a_1)$, dann wird ausgehend vom Intervall (a_1, a_2) , ein äußerer Punkt a_3 bestimmt, so daß a_2 die Strecke $\overline{a_1 a_3}$ harmonisch teilt. Danach ist zu untersuchen, ob im neuen Intervall ein Minimum enthalten ist ($f(a_3) < f(a_2)$?) und so fort.

Die Aufweitung ist abzubrechen, wenn

$$f(a_{n+1}) > f(a_n)$$

ist.

Die Intervallbreite bei der Aufweitung wächst in geometrischer Progression. Zwei aufeinanderfolgende Intervalle verhalten sich wie $(1 + \sqrt{5})/2$. Der n -te Punkt

hat die Abszisse

$$a_n = \frac{\sqrt{5}-1}{2} \left[\left(\frac{\sqrt{5}+1}{2} \right)^n - 1 \right]. \quad (37)$$

Für den dritten Teil der eindimensionalen Minimumsuche, die Interpolation, bietet sich ein Polynom dritten Grades an. Es läßt sich aus den vier Punkten des letzten Schachtelintervalls bestimmen und besitzt ein eindeutiges Minimum. Bei der Bestimmung dieses Minimums ist zu beachten, daß das Polynom dritten Grades zu einer Parabel entarten kann.

Sollte der Funktionswert an der interpolierten Stelle größer sein als der Funktionswert zu Beginn der eindimensionalen Minimumsuche (dies kann bei extrem nichtlinearen Funktionen geschehen), dann ist die Schachtelung fortzusetzen. Ergibt sich auch dabei kein kleinerer Funktionswert, dann ist die Iteration in ein höher gelegenes Nachbartal gerutscht.

In Abschnitt 5 wurde gesagt, in scheinbarem Widerspruch zu diesen Ausführungen, für die Iteration sei eine genaue Minimumbestimmung nicht erforderlich. Das hier beschriebene aufwendige Verfahren soll lediglich gewährleisten, daß die Funktion bei jedem Iterationsschritt abnimmt. Einfachere Methoden führen bei komplizierten Funktionen zur Divergenz.

Im verwendeten Programm wird nach der Intervallaufweitung gleich interpoliert. Die Schachtelung setzt nur ein, wenn der Funktionswert an der interpolierten Stelle nicht kleiner geworden ist.

8. Beispiele

Das beschriebene Verfahren wurde programmiert und an Funktionen mit maximal 50 Variablen erprobt. In einem Programm zur Optimierung linearer elektrischer Netzwerke bewährt es sich seit über einem Jahr. Für die Neben-

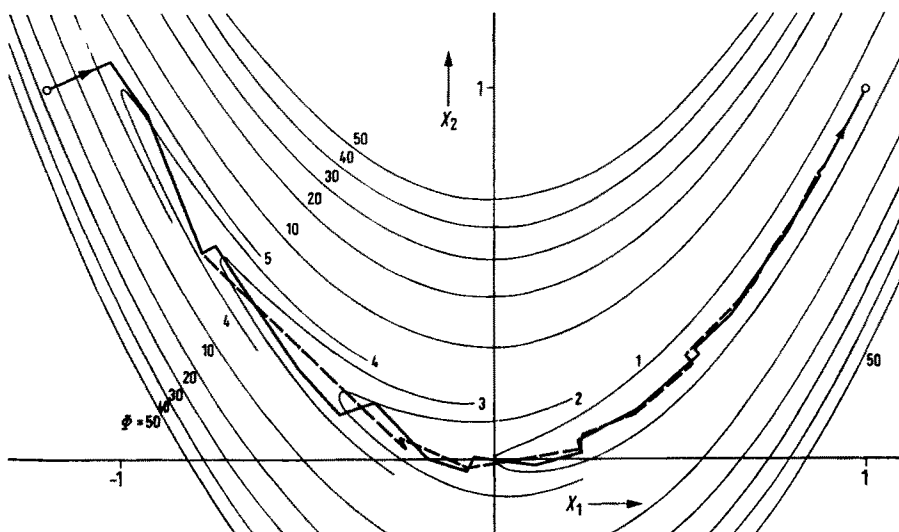


Fig. 4. Iteration an der Funktion $\Phi = (1-x_1)^2 + 100(x_2 - x_1^2)^2$ (Bananafunktion).
— Neue Methode. --- Fletcher-Powell

bedingungen wurde dabei nicht die allgemeine Form (2) vorgesehen sondern die einfachere

$$x_{i\min} \leq x_i \leq x_{i\max}.$$

Weil es in diesem speziellen Fall bequemer ist, wurden die Eigenvektoren nicht aus $|H|$, sondern aus der Matrix $|H^{-1}|$ gewonnen. $|H^{-1}|$ wurde in gleicher Weise wie $|H|$ aufgebaut; im Algorithmus sind dabei nur $|\Delta g\rangle$ und $|\Delta x\rangle$ zu vertauschen. Zur Berechnung der benötigten Eigenwerte und Eigenvektoren diente ein abgewandeltes Hessenberg-Verfahren [5].

Die folgenden Beispiele sollen einen Eindruck von der Arbeitsweise des Verfahrens vermitteln. In der zum Vergleich angegebenen Fletcher-Powell-Methode wurde dasselbe oben beschriebene Unterprogramm für die eindimensionale Minimumsuche benutzt. Vergleichen lassen sich die beiden Verfahren natürlich nur für Iterationen ohne Nebenbedingungen.

Fig. 4 zeigt die Minimierung der Funktion $\Phi = (1 - x_1)^2 + 100(x_2 - x_1^2)^2$. Diese Funktion (Bananafunktion) wurde von Rosenbrock in die Minimierungsliteratur eingeführt und zeichnet sich durch eine parabelförmige, extrem schmale Rinne aus, durch die die Iteration sich durchschlängeln muß. Die neue Methode benötigt vom Startpunkt $|-1.2, 1\rangle$ zur Lösung $|1, 1\rangle$ 25 Schritte, die nach Fletcher und Powell 18.

Bei dem in Fig. 5 gezeigten Beispiel wird eine Funktion mit Sattelpunkt minimiert: $\Phi = \frac{1}{4}x_1 + (x_1^2 - 2x_1 + x_2^2)^2$. Die Funktionsfläche hat die Gestalt eines schräg gestellten Hutes mit ausladender Krempe (Sombrero). Die vom Startpunkt zur Lösung führende Rinne ist nicht so extrem schmal wie im obigen Beispiel. Hier wurde nicht mit einem Gradientenschritt, sondern mit einem Matrixaufbau begonnen. Die Matrix der zweiten Ableitungen besitzt in der Nähe des

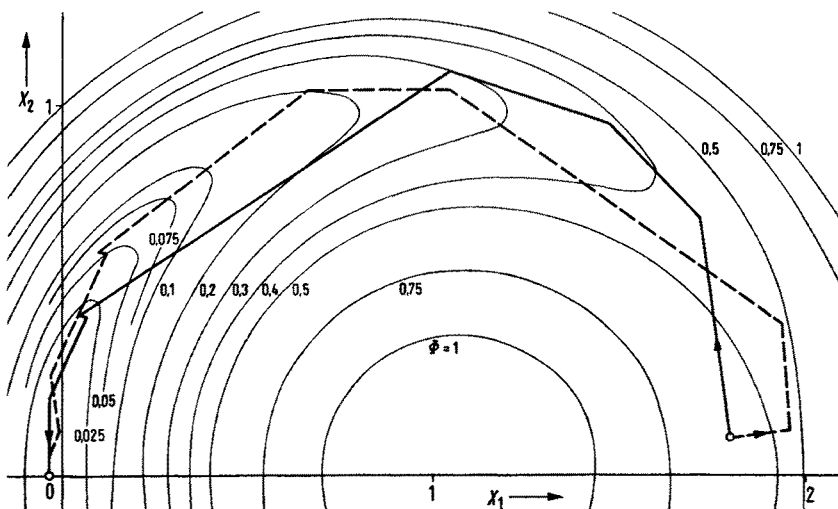


Fig. 5. Iteration an der Funktion $\Phi = \frac{1}{4}x_1 + (x_1^2 - 2x_1 + x_2^2)^2$. — Neue Methode.
--- Fletcher-Powell

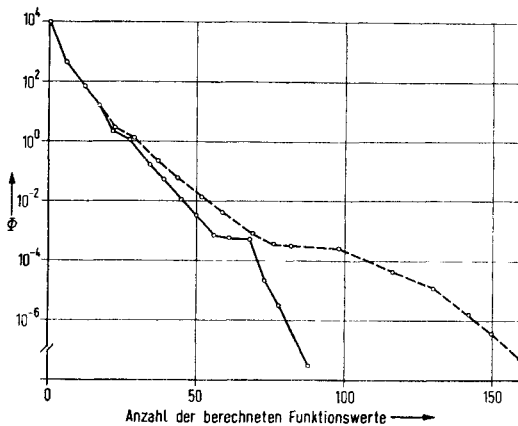


Fig. 6

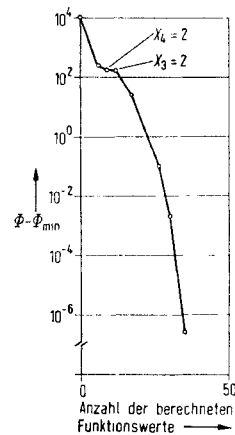


Fig. 7

Fig. 6. Iteration ohne Nebenbedingungen an der Funktion
 $\Phi = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$.
 — Neue Methode. --- Fletcher-Powell

Fig. 7. Iteration unter den Nebenbedingungen $x_3 \geq 2$, $x_4 \geq 2$ an der Funktion
 $\Phi = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$

Sattels einen negativen Eigenwert, und der erste Schritt führt parallel zur Talsohle vom Sattel weg. Die neue Methode benötigt hier 7, die von Fletcher und Powell 11 Schritte.

In Fig. 6 ist der Funktionswert über der Anzahl der berechneten Funktionswerte beim Minimieren der Funktion

$$\Phi = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$$

aufgetragen. Diese Funktion wurde von Fletcher und Powell angegeben. Im Minimum ($|x\rangle = |0, 0, 0, 0\rangle$) ist die Matrix der zweiten Ableitungen dieser Funktion doppelt singulär. Gestartet wurde in dem Punkt $|x\rangle = |1, 2, 3, 4\rangle$. An dieser Funktion zeigt das neue Verfahren gegenüber dem Fletcher-Powellschen deutlich Vorteile. Es konvergierte von allen angesetzten Startpunkten aus wesentlich schneller.

Eine Iteration unter Nebenbedingungen zeigt Fig. 7. Funktion und Startvektor sind die gleichen wie in Fig. 6. Nebenbedingungen sind die beiden Ungleichungen $x_3 \geq 2$ und $x_4 \geq 2$. Die Iteration ist hier schneller am Ziel ($|x\rangle = |1,27 \dots, 0,63 \dots, 2, 2\rangle$), weil nach Erreichen der Begrenzungsebenen effektiv nur mehr ein zweidimensionaler Raum der Variation zur Verfügung steht. Ein zweidimensionales Problem bereitet weniger Mühe als ein vierdimensionales.

Für die folgenden Testbeispiele mußte die Strategie des Verfahrens abgeändert werden. Die Funktionen erwiesen sich als so ausgeprägt nichtlinear, daß sich mit fest vorgegebenen Probestepbewegungen beim Aufbau der Iterationsmatrix nach einem Schritt längs des Eigenvektors keine brauchbaren Näherungsmatrizen er-

gaben. Dieser spezielle Matrixaufbau wurde deshalb fallen gelassen. Die Funktionen waren von der Form

$$\Phi = \sum_{k=1}^1 \left[c \frac{\prod_{i=1}^m (1 + z_{1i} x_k + z_{2i} x_k^2)}{\prod_{j=1}^m (1 + n_{1j} x_k + n_{2j} x_k^2)} - f_k \right]^2.$$

Die Sollwerte f_k wurden aus der Beziehung

$$f(x) = c \frac{\prod_{i=1}^m [(x+1-3i)^2 + \frac{1}{2}]}{\prod_{i=1}^m [(x+\frac{1}{2}-3i)^2 + \frac{1}{2}]}$$

gewonnen, wobei c gerade so gewählt war, daß f für $x=0$ den Wert 1 hatte. Variiert wurden die Parameter z_{1i} , z_{2i} , n_{1j} , n_{2j} und c . Als Ausgangsnäherung wurden einfachheitshalber für die Parameter z und n die exakten Lösungswerte eingegeben, nur der Parameter c war von der Lösung verschieden. Bei der Iteration bewegten sich die Parameter z und n zunächst von den Lösungen fort und kehrten dann wieder zu ihnen zurück, ähnlich dem Parameter x_2 bei der „Bananafunction“.

In den Fig. 8 und 9 sind die Iterationsabläufe für zwei Beispiele aufgezeichnet. Dabei ist der Funktionswert in logarithmischem Maßstab über der Anzahl der berechneten Funktionswerte aufgetragen. Zum Vergleich ist die Iteration nach der Methode von Fletcher und Powell gestrichelt eingezeichnet. Für $m=3$ erhält man eine 13-Parameter-Funktion. Fig. 8 zeigt die Iteration an dieser Funktion, wobei $l=24$ Stützpunkte im Bereich $0,5 \leq x \leq 12,0$ vorgegeben waren. Die neue Methode rechnete 61 Schritte und benötigte dazu 753 Funktionswerte. Bei

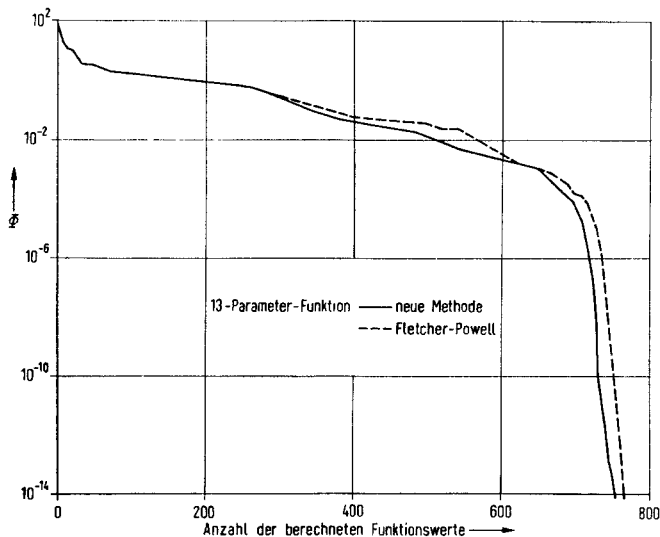


Fig. 8. Iteration an einer 13-Parameter-Funktion

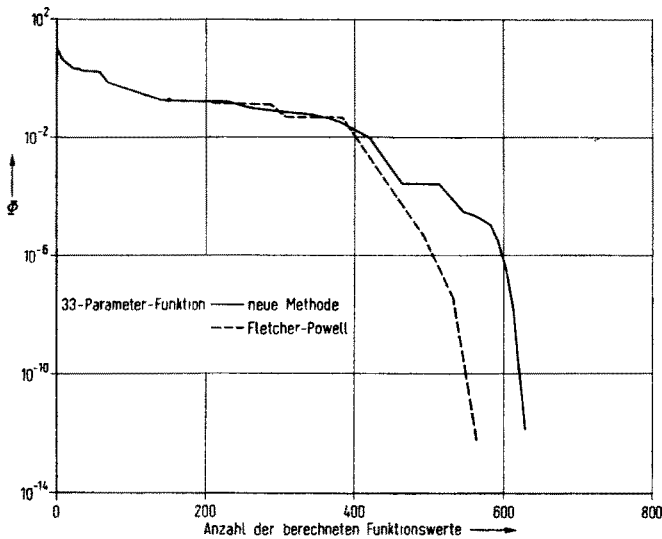


Fig. 9. Iteration an einer 33-Parameter-Funktion

Fletcher und Powell waren es 69 Schritte und 766 Funktionswerte. Setzt man $m=8$, dann besitzt die Funktion 33 Parameter. In Fig. 9 ist die Minimierung dieser Funktion für $l=52$ Stützwerte im Bereich $0,5 \leq x \leq 26,0$ dargestellt. Beide Verfahren rechnen 57 Schritte. Die neue Methode berechnet dazu die Funktion 629mal, Fletcher und Powells Verfahren 584mal.

Literatur

1. Bellmann, R.: Dynamic programming. Princeton: Princeton University Press 1962.
2. Dirac, P. A. M.: The principles of quantum mechanics. Oxford: Clarendon Press 1958.
3. Fiacco, A. V., McCormick, G. P.: The sequential unconstrained minimization technique for nonlinear programming, a primal-dual method. *Management Science* **10**, 360—366 (1964).
4. — — Computational algorithm for the sequential unconstrained minimization technique for nonlinear programming. *Management Science* **10**, 601—617 (1964).
5. Fletcher, R., Powell, M. J.: A rapidly convergent descent method for minimization. *The Computer Journal* **6**, 163 (1963).
6. Lasdon, L. S., Waren, A. D.: Optimal design of filters with bounded, lossy elements. *IEEE Transactions CT-13* (June 1966), S. 175—187.
7. Zurmühl, R.: *Matrizen*, S. 321 ff. Berlin-Göttingen-Heidelberg: Springer 1964.

G. Zettl
Siemens AG
D-8 München 25
Postfach