Kapitel 1I

Das Verfahren von Hildreth und d'Esopo

1. Dualisierung des Verfahrens

Man geht aus vom Problem (*PIII*):

$$Min\left\{Q(x) = x^T Cx + p^T x \mid Ax \le b\right\},$$

aus, wobei A eine $(m \cdot n)-$Matrix und C eine positiv definite, symmetrische Matrix der Ordnung n ist.[1]

Auf Grund der strengen Konvexität der Zielfunktion hat (*PIII*), sofern es zulässig ist, eine eindeutige Lösung x^*[2].

Die Kuhn-Tucker-Bedingungen für das Problem (*PIII*) lassen sich (siehe Kapitel I) wie folgt schreiben:

$$(1.\ 10.)\qquad \begin{cases} Ax + y = b & (A) \\ 2Cx + A^T u = -p & (B) \\ u \ge 0,\ y \ge 0 & (C) \\ u^T y = 0 & (D) \end{cases}$$

Da x eine unbeschränkte Variable ist und C^{-1} existiert[3], kann man x von diesen Bedingungen eliminieren. Aus der Beziehung (1. 10(B)) folgt

$$(2.\ 1.)\qquad x = -\frac{1}{2}C^{-1}(A^T u + p).$$

Substituiert man (2. 1.) in (1. 10 (A)), so erhält man

$$-\frac{1}{2}AC^{-1}(A^T u + p) + y = b,$$

d. h.

$$-\frac{1}{2}AC^{-1}A^T u + y = b + \frac{1}{2}AC^{-1}p$$

---

[1] Es wird weiter vorausgesetzt (siehe den Konvergenzbeweis), dass für die schärferen Restriktionen

$$a_i^T x \le b_i - \varepsilon \ (i = 1,2,...,m)$$

auch noch zulässige Punkte existieren. Hierbei ist a_i^T die $i-$te Zeile von A, b_i die $i-$te Komponente von b und ε eine beliebig kleine positive Größe.

[2] **Satz**: Eine streng konvexe Funktion hat höchstens ein Minimum.

[3] **Satz**: Eine streng definite Matrix ist nichtsingulär.

$$2\left(\frac{1}{4}AC^{-1}A^T\right)u - y = -\left(b + \frac{1}{2}AC^{-1}p\right)$$

bzw.

(2. 2.) $$2GU - y = -h,$$

wobei

$$G = \frac{1}{4}AC^{-1}A^T,$$
$$h = b + \frac{1}{2}AC^{-1}p.$$

Damit werden die Kuhn-Tucker-Bedingungen auf die folgende Form reduziert:

(1. 14.) $$\begin{cases} 2Gu - y = -h \\ u \geq 0,\ y \geq 0 \\ u^T y = 0. \end{cases}$$

Die obigen Bedingungen sind aber, wie im Kapitel I gezeigt wurde, die Kuhn-Tucker-Bedingungen für das Problem

(2. 3.) $$Min\left\{\Theta(u) = u^T Gu + h^T u \mid u \geq 0\right\},$$

welches das Dualproblem (*D*III*) zum Problem (*PIII*) ist. Hierbei sind zwei Fälle zu unterscheiden:

a) Die Matrix G ist positiv definit, wenn A den vollen Zeilenrang hat[4], was besonders in der Ökonomie (infolge der Anforderung der Beschränktheit des zulässigen Bereiches) nicht häufig der Fall sein wird. Wir werden diesen Fall als Problem (*PIIIa*) bezeichnen.

<u>Satz 1</u>:
Problem (*PIIIa*) hat eine eindeutige Lösung.

B e w e i s:
Es sei u^* diese Lösung. Ihre Existenz ist gesichert dadurch, dass $\Theta(u)$ eine untere Schranke hat und die Menge der zulässigen u abgeschlossen ist.
Nehmen wir an, dass $\Theta(u)$ in mindestens zwei Punkten u^* und u^{**} $(u^* \neq u^{**})$ ein Minimumbesitzt. Aus der Konvexität von $\Theta(u)$ folgt für $0 < \alpha < 1$:

(2. 4.) $$\Theta\left\{\alpha u^* + (1-\alpha)u^{**}\right\} < \alpha\Theta(u^*) + (1-\alpha)\Theta(u^{**})$$

[4] **<u>Satz</u>**: Es sei C eine streng definite, symmetrische $(n \cdot n)$ – Matrix . Die symmetrische $(m \cdot m)$ – Matrix $AC^{-1}A^T$ ist positiv definit, wenn A den vollen Zeilenrang $r = m$ hat. Andernfalls ist sie positiv semidefinit.

$$= \Theta(u^*) = \Theta(u^{**}),$$

d.h.

(2. 5.) $$\Theta\{\alpha u^* + (1-\alpha)u^{**}\} < \Theta(u^*).$$

Das ist aber im Widerspruch zu unserer Voraussetzung, dass u^* und u^{**} zugleich die minimierenden Lösungen des Problems (*PIIIa*) sind. Damit ist der Beweis beendet.

b) Die Matrix G ist positiv semidefinit, wenn die Matrix A nicht den vollen Zeilenrang hat. Im Weiteren werden wir diesen, von der ökonomischen Praxis aus gesehen, interessanteren Fall mehr Aufmerksamkeit schenken:
Die Matrix G hat auf der Hauptdiagonale positive Elemente, d.h.

(2. 6.) $$g_{ii} = \frac{1}{4} a_i^T C^{-1} a_i > 0,$$

Weil die Vektoren a_i verschieden von Null sind und die Inverse C^{-1} positiv semidefinit ist.[5]
Aus dem Angeführten folgt, dass x^* die optimale Lösung des Problems (*PIII*) darstellt dann und nur dann wenn

$$x^* = -\frac{1}{2} C^{-1}(A^T u^* + p)$$

Gilt, wobei u^* *eine* optimale Lösung des Problems (2. 3.) ist.
Es ist zu bemerken, dass das Dualproblem (2. 3.) im Vergleich zum Primalproblems (*PIII*) einfacher ist, da seine Restriktionen nur aus den Nichtnegativitätsbedingungen bestehen.

2. Lösung des Dualproblems

Nach dem Satz 4von Kapitel I, wenn das Primalproblem (*PIII*) lösbar ist, so hat auch das Dualproblem (2. 3.) eine Lösung. Wenn die Lösung von (2. 3.) nicht eindeutig ist, so ergeben alle optimalen Lösungen u^* das gleiche

$$x^* = -\frac{1}{2} C^{-1}(A^T u^* + p) = x(u^*).$$

Zur Auflösung des Dualproblems kann man die folgende Methode (eine Abwandlung des Gauß-Seidelschen Iterationsverfahrens [82]) anwenden.
Als Anfangslösung wählen wir einen beliebigen Vektor $u^0 \geq 0$. (Es ist im Allgemeinen günstig, $u^0 = 0$ zu setzen, da einige Komponenten des optimalen Lösungsvektors den Wert Null haben). Die Komponenten des nächsten (verbesserten) Iterationspunktes u^1 erhalten wir, indem wir die Funktion $\Theta(u) = u^T GU + h^T u$ (bezüglich der Menge $u \geq 0$) der Reihe nach hinsichtlich der i – ten Komponente des Vektors u minimieren, wobei die anderen Komponenten u_j $(j \neq i)$ den zuletzt fixierten Wert innehaben, d.h.

[5] **Satz:** Ist C positiv (negativ) definit, so ist auch die Inverse C^{-1} positiv (negativ) definit.

für $j > i \quad u_j = u_j^0$

und

für $j < i \quad u_j = u_j^1$.

Das Minimum der Funktion $\Theta(u_i) = (u_1^1, u_2^1, ..., u_{i-1}^1, u_i, u_{i+1}^1, ..., u_m^1)$ wird folgendermaßen bestimmt:
Zuerst ermitteln wir mit Hilfe der ersten partiellen Ableitung das freie Minimum der Funktion $\Theta(u_i)$, welches wir als w_i^1 bezeichnen.
Auf Grund der Beziehung

$$\frac{\partial \Theta(u)}{\partial u_i} = 2\sum_{j=1}^{m} g_{ij} u_j + h_i$$

gilt

$$0 = \frac{\partial \Theta(u_i)}{\partial u_i} = 2\sum_{j=1}^{i-1} g_{ij} u_j^1 + 2\sum_{j=i+1}^{m} g_{ij} u_j^0 + h_i.$$

Daraus folgt

$$w_i^1 = -\frac{1}{g_{ii}} \left\{ 2\sum_{j=1}^{i-1} g_{ij} u_j^1 + 2\sum_{j=i+1}^{m} g_{ij} u_j^0 + \frac{1}{2} h_i \right\}.$$

Für $w_i^1 \geq 0$ gilt $u_i^1 = w_i^1$. Ist dagegen $w_i^1 < 0$, d.h. liegt das freie Minimum der Funktion $\Theta(u_i)$ außerhalb des Bereiches $u \geq 0$, so setzt man $u_i^1 = 0$. In diesem Falle stellt das Minimum der Funktion $\Theta(u_i)$ bezüglich $u_i \geq 0$ einen „Randpunkt“ dar. Die Regel lässt sich formell folgendermaßen schreiben:

$$u_i^1 = \max(0;\ w_i^1).$$

Allgemein findet man u^{k+1} für $k = 0, 1, 2, ...$ durch die Formel

(2. 7.) $$u_i^{k+1} = \max(0;\ w_i^{k+1}) \text{ für } k = 0,\ 1,\ 2, ...$$

mit

(2. 8.) $$w_i^1 = -\frac{1}{g_{ii}} \left\{ 2\sum_{j=1}^{i-1} g_{ij} u_j^{k+1} + 2\sum_{j=i+1}^{m} g_{ij} u_j^k + \frac{1}{2} h_i \right\},\ i = 1,\ 2, ...,\ m.$$

Offensichtlich gilt

(2. 9.) $$\Theta(u^{k+1}) \leq \Theta(u^k).$$

Im nächsten Abschnitt werden wir beweisen, dass die Folge $\{\Theta(u^k)\}$ gegen den Minimalwert $\Theta(u^*)$ und die Folge $\{x(u^k)\}$ gegen die optimale Lösung $x^* = x(u^*)$ konvergiert.

Die Iterationen müssen solange fortgeführt werden, bis sich $\Theta(u^k)$ bzw. $x(u^k)$ im Rahmen der gewünschten Rechengenauigkeit nicht mehr ändert.

3. Konvergenzbeweis

Es muss gezeigt werden, dass

$$\lim_{k\to\infty}\Theta(u^k)=\Theta(u^*).$$

Wir schreiben symbolisch $u^{k+1}=P(u^k)$, wobei man $P(u)$ dadurch definiert, dass man (2. 7.) und (2. 7.) der Reihe nach für $i=1,2,...,m$ anwendet, also etwas ausgeschrieben

a) $v^0=u$

b) $$\begin{cases} v_j^{\upsilon+1}=v_j^{\nu} \text{ für } j\neq\upsilon+1 \\ v_{\upsilon+1}^{\upsilon+1}=\max\left\{0;\ -\dfrac{1}{g_{\upsilon+1,\upsilon+1}}\left(\sum\limits_{p\neq\upsilon+1} g_{p,\upsilon+1}v_p^{\upsilon}+\frac{1}{2}h\upsilon+1\right)\right\} \end{cases}$$

c) $P'(u)=v^m$.

(v^{υ} stimmt in υ Komponenten mit hu und in $m-\upsilon$ Komponenten mit $P(u)$ überein).

Für festes υ hängen alle $v_j^{\upsilon+1}$ stetig von v_j^{υ} ab. Das ist trivial für diejenigen Komponenten, die sich überhaupt nicht ändern, und wird auch für die restliche Komponente leicht verifiziert. Da $v_j^{\upsilon+1}$ stetig von v_j^{υ} abhängt, hängt auch v^m stetig von v^0 ab (Kopplung stetiger Abbildungen). Damit ist die Stetigkeit von $P(u)$ bewiesen.
Es wird nun der Beweis erbracht, dass analog zu (2. 9.) gilt:

(2. 10.) $$\Theta\big[P(u)\big]\le\Theta(u),$$

wobei (2. 10.) in Gleichungsform, d.h.

(2. 11.) $$\Theta\big[P(u)\big]=\Theta(u)$$

gilt, dann und nur dann, wenn $u=u^*$ ist.

Beim Übergang von v^{υ} zu $v^{\upsilon+1}$ ändert sich nur eine Komponente und diese wird so bestimmt, dass die Funktion

(2. 12.) $$\Phi(\lambda)=\Theta(v_1^{\upsilon},...,v_{\upsilon}^{\upsilon},\lambda,v_{\upsilon+2}^{\upsilon},...)$$

$$=\Theta(v_1^{\upsilon+1},...,v_{\upsilon}^{\upsilon+1},\lambda,v_{\upsilon+2}^{\upsilon+1})$$

ihr Minimum unter der Bedingung $\lambda\ge 0$ im Punkt $\lambda=v_{\upsilon+1}^{\upsilon+1}$ annimmt (das ist der Sinn von (2. 8.)). Es gilt also

$$\Phi\left(v_{\upsilon+1}^{\nu+1}\right) \le \Phi\left(v_{\upsilon+1}^{\nu}\right),$$

d.h.

$$\Theta\left(v^{\upsilon+1}\right) \le \Theta\left(v^{\upsilon}\right)$$

und damit

$$\Theta\left(v^{m}\right) \le \Theta\left(v^{0}\right),$$

d.h.

$$\Theta\left[P(u)\right] \le \Theta(u).$$

Wegen $g_{\upsilon+1,\upsilon+1} > 0$ ist $\Phi(\lambda)$ ist eine streng konvexe quadratische Funktion in λ; das Minimum von $\Phi(\lambda)$ unter der Bedingung $\lambda \ge 0$ ist nicht eindeutig (das gilt nebenbei nicht für $\Theta(u)$; $\Theta(u)$ kann im Bereich $u \ge 0$ mehrere Minima haben. Anders ausgedrückt: $\Theta(u)$ ist streng konvex längs jeder Koordinate, aber nicht streng konvex in allen Koordinaten zusammen).
$\Theta\left[P(u)\right] = \Theta(u)$ ist nur dann möglich, wenn

$$\Theta\left(v^{\upsilon+1}\right) = \Theta\left(v^{\upsilon}\right), \quad \upsilon = 0,1,...,m-1.$$

Wegen der strengen Konvexität von $\Phi(\lambda)$ und der damit verbundenen Eindeutigkeit des Minimum folgt aus $\Theta\left(v^{\upsilon+1}\right) = \Theta\left(v^{\upsilon}\right)$, dass

$$v^{\upsilon+1} = v^{\upsilon},$$

d.h. aus $\Theta\left[P(u)\right] = \Theta(u)$ folgt

$$P(u) = u.$$

Dies bedeutet aber, dass $\Theta(u)$ in u das Minimum parallel zu jeder Koordinatenachse annimmt, d.h. wenn man alle Komponenten bis auf eine festhält, so lohnt es sich auch nicht, diese eine auch abzuändern. Man hat dann

$$\text{(2. 13.)} \qquad \begin{cases} \dfrac{\partial\Theta}{\partial u_j} \ge 0 \\ u_j \ge 0 \\ \dfrac{\partial\Theta}{\partial u_j} \cdot u_j = 0 \end{cases} \qquad \text{für alle } j.$$

(Wird dies für die j – te Komponente nicht erfüllt, so könnte man eben Θ durch Änderung der j – te Komponente noch verringern).
Die Bedingungen (2. 13.) sind aber die Kuhn-Tucker-Bedingungen für das Problem (2. 3.) und damit ist die Beziehung (2. 11.) für $u = u^*$ bewiesen.

Es wird nun gezeigt, dass der Bereich

(2. 14.) $\left\{u \mid u \geq 0,\ \Theta(u) \leq \Theta(u^0)\right\}$

Beschränkt ist:
Wäre dies nicht der Fall, so könnte man die Funktion

(2. 15.) $\Theta_\varepsilon(u) = \Theta(u) - \varepsilon \sum_{i=1}^{m} u_i$

Unter der Nichtnegativitätsbedingung $u \geq 0$ *ad infinitum* minimieren. (2. 15.) ist aber die duale Zielfunktion zu dem modifizierten Primalproblem (vgl. Fußnote 1):

(*PIII'*) $Min\left\{Q(x) = x^T CX + p^T x \mid a_i^T x \leq b_i - \varepsilon,\ i = 1,2,...,m\right\}.$

Auf Grund der Voraussetzung über den zulässigen Bereich von (*PIII*) hat auch (*PIII'*) eine Lösung, und nach dem Dualitätsprinzip muss $\Theta_\varepsilon(u)$ auf der Menge $u \geq 0$ beschränkt sein. Durch diesen Widerspruch ist die Beschränktheit von (2. 14.) bewiesen.

Wegen der Beschränktheit von (2. 14.) besitzt die Menge der Iterationslösungen u^k mindestens einen Häufungspunkt $\bar{u}$ (Bolzano -Weierstraß-Satz). Auf Grund der Stetigkeit von $\Theta(u)$ ist $\Theta(\bar{u})$ Häufungspunkt der Folge $\Theta(u^k)$. Wegen der Stetigkeit von $P(u)$ ist $P(\bar{u})$ Häufungspunkt von $P(u^k)$, und damit ist auch $\Theta\left[P(\bar{u})\right]$ Häufungspunkt der Folge $\Theta\left[P(u^k)\right] = P(u^{k+1})$, und daher zugleich der Folge $\Theta(u^k)$. Da die Folge $\Theta(u^k)$ abnimmt, müssen die beiden Häufungspunkte übereinstimmen:

(2. 16.) $\Theta(\bar{u}) = \Theta[P(\bar{u})],$

d.h.

(2. 17.) $\bar{u} = u^*.$

Da eine abnehmende Folge gegen ihren Häufungspunkt konvergiert gilt

(2. 18.) $\lim_{k\to\infty} \Theta(u^k) = \Theta(u^*).$

Damit ist der Beweis beendet.

4. Rechenschema, Flussdiagramm und ein Illustrationsbeispiel

Das Rechenschema für die Berechnung von u_i^{k+1} bzw. w_i^{k+1}, $k = 0,1,2,...$, gemäß (2. 7.) und (2. 8.) ist einfach zu konstruieren. Es sieht für den Fall $m = 4$ folgendermaßen aus:

Rechenschema des Algorithmus von Hildreth – d´Esopo

k					
	g_{11}	g_{12}	g_{13}	g_{14}	$\frac{1}{2}h_1$
	g_{21}	g_{22}	g_{23}	g_{24}	$\frac{1}{2}h_2$
	g_{31}	g_{32}	g_{33}	g_{34}	$\frac{1}{2}h_3$
	g_{41}	g_{42}	g_{43}	g_{44}	$\frac{1}{2}h_4$
0	$u_1^0 = 0$	$u_2^0 = 0$	$u_3^0 = 0$	$u_4^0 = 0$	1
1	u_1^1	u_2^1	u_3^1	u_4^1	1
2	u_1^2	u_2^2	u_3^2	u_4^2	1
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
$k-1$	u_1^{k-1}	u_2^{k-1}	u_3^{k-1}	u_4^{k-1}	1
k	u_1^k	u_2^k	u_3^k	u_4^k	1
$k+1$	u_1^{k+1}	u_2^{k+1}	u_3^{k+1}	u_4^{k+1}	1
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

Koeffiziententableau

Iterationstableau

Das *Iterationstableau* wird Schritt für Schritt ausgeführt. Der Algorithmus wird an der Berechnung von u_3^{k+1} illustriert:

Rechenvorschrift:

1. Man bilde für $j < 3$ das Skalarprodukt von den Komponenten der $(k+1)-$ten Iteration und den entsprechenden Elementen der dritten Zeile im Koeffizienten-tableau.

2. Man berechne für $j > 3$ das Skalarprodukt von den Komponenten der $k-$ten Iteration und den entsprechenden Elementen der dritten Zeile, einschließlich der letzten Spalte im Koeffiziententableau.

3. Man addiere die Skalarprodukte der in 1. und 2. Zusammen und dividiere das Resultat durch das negative Diagonalelement g_{33}.

4. Falls das Ergebnis w_3^{k+1} positiv ist, so wird es als u_3^{k+1} eingetragen.

5. Falls das Ergebnis w_3^{k+1} nicht positiv ist, so setzt man $u_3^{k+1} = 0$.

Zusammengefasst:

$$u_3^{k+1} = \max\left(w_3^{k+1};\ 0\right),$$

$$w_3^{k+1} = -\frac{1}{g_{33}}(g_{31}u_1^{k+1} + g_{32}u_2^{k+1} + g_{32}u_4^k + \frac{1}{2}h_3).$$

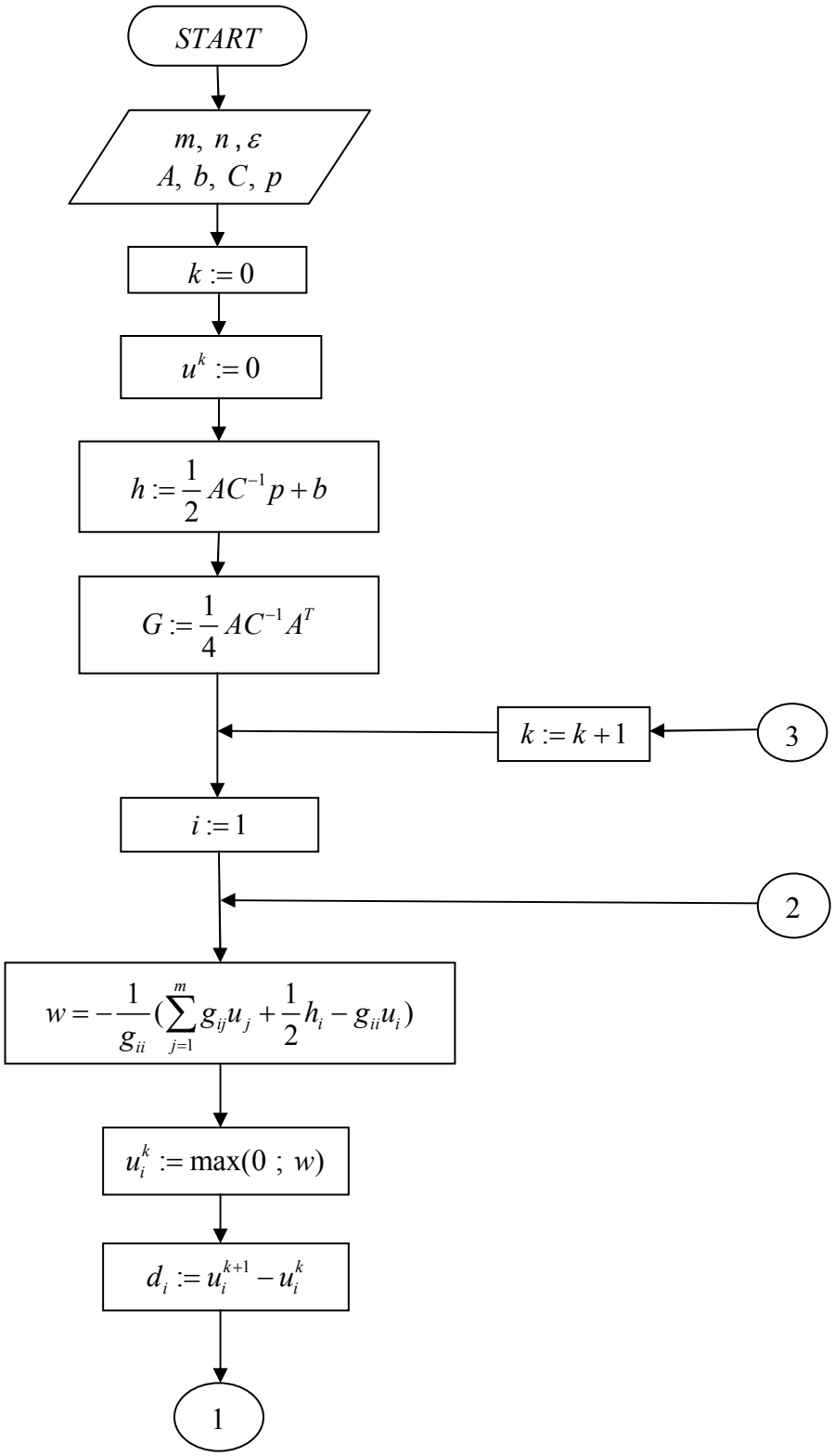
Anmerkung:

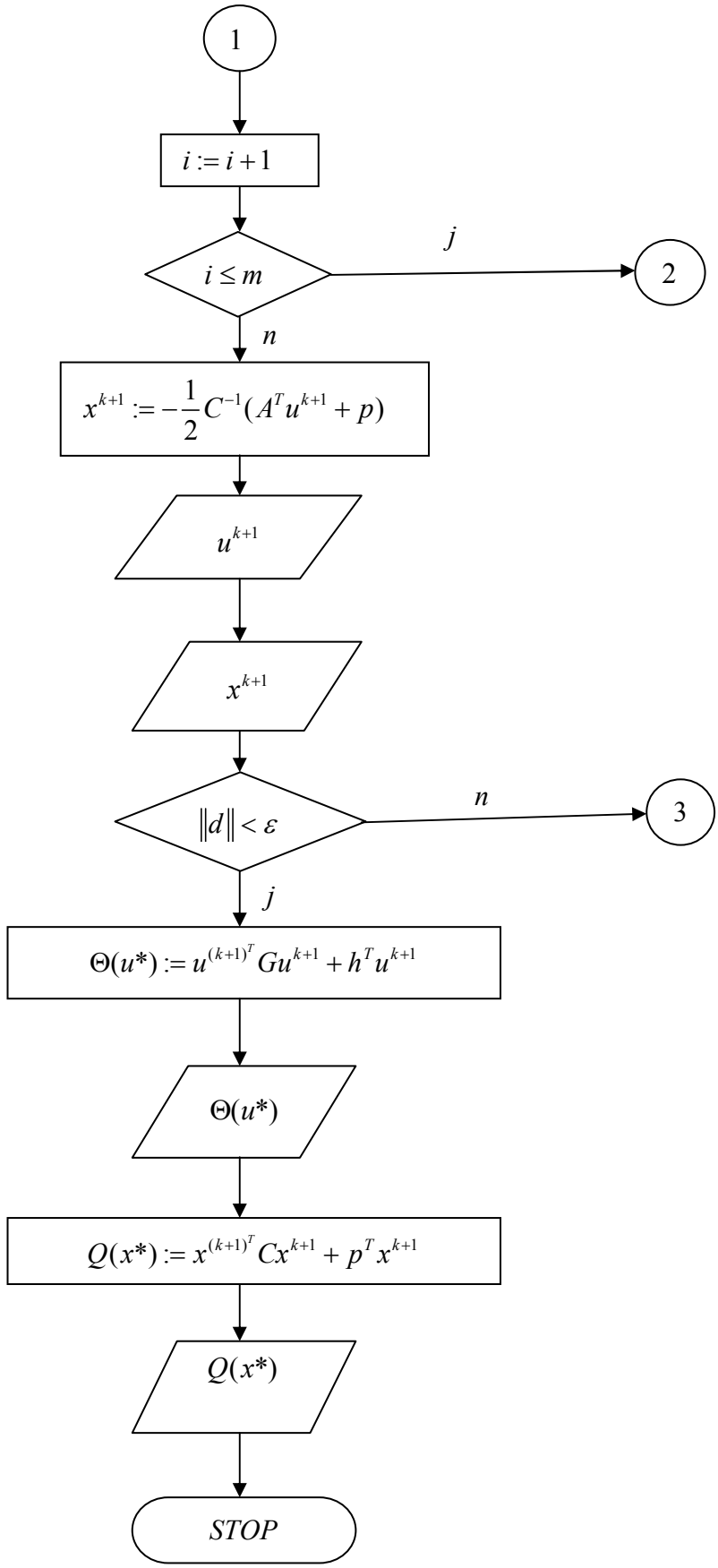
1. Multipliziert man das Koeffiziententableau mit einer positiven Zahl, so erhält man dieselbe Iterationsfolge. Diese Eigenschaft ermöglicht uns manchmal, die Berechnung mit ganzen Zahlen durchzuführen.

2. Der Iterationsprozess kann mit beliebiger Komponente des Vektors u^{k+1} beginnen. Natürlich erhält man dann eine andere Iterationsfolge, die aber auch gegen u^* konvergiert. Die Anzahl der benötigten Schritte kann sich jedoch nicht ändern.

3. Eventuelle Rechenfehler werden keinen Einfluss auf das Resultat haben. Sie können nur die Konvergenz beeinflussen.

Flussdiagramm des Verfahrens von Hildreth und d'Esopo




1
$i := i + 1$
$i \le m$
j
2
n
$x^{k+1} := -\frac{1}{2} C^{-1}(A^T u^{k+1} + p)$
u^{k+1}
x^{k+1}
$\|d\| < \varepsilon$
n
3
j
$\Theta(u^*) := u^{(k+1)^T} G u^{k+1} + h^T u^{k+1}$
$\Theta(u^*)$
$Q(x^*) := x^{(k+1)^T} C x^{k+1} + p^T x^{k+1}$
$Q(x^*)$
STOP

Illustrationsbeispiel:

$$\min Q(x) = 2x_1^2 - 6x_1x_2 + 5x_2^2 - 3x_1 - x_2$$

unter den Restriktionen

$$\begin{aligned} 2x_1 \quad - x_2 &\le 7 \\ x_1 + 2x_2 &\le 10 \\ x_1 \qquad &\ge 0 \\ x_2 &\ge 0. \end{aligned}$$

In unserer Schreibweise ist

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad b = \begin{pmatrix} 7 \\ 10 \\ 0 \\ 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}, \qquad p = \begin{pmatrix} -3 \\ -1 \end{pmatrix}.$$

Man berechnet zuerst

$$C^{-1} = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$$

und hierauf G und h :

Demonstrationsbeispiel

	10 15 -7 -4	15 25 -11 -7	-7 -11 5 3	-4 -7 3 2	-11 -20 18 11
u^0	0.000	0.000	0.000	0.000	1
u^1	1.100	0.140	0.000	0.000	1
u^2	0.890	0.270	0.000	0.000	1
u^3	0.700	0.380	0.000	0.000	1
u^4	0.530	0.480	0.000	0.000	1
u^5	0.380	0.570	0.000	0.000	1
u^6	0.240	0.660	0.000	0.000	1
u^7	0.110	0.730	0.000	0.000	1
u^8	0.005	0.787	0.000	0.000	1
u^9	0.000	0.800	0.000	0.000	1
u^{10}	0.000	0.800	0.000	0.000	1

Daher ist

$$u^* = \begin{pmatrix} 0 & 4/5 & 0 & 0 \end{pmatrix}^T$$

und

$$x^* = -\frac{1}{2} C^{-1}(A^T u^* + p) = \begin{pmatrix} \frac{23}{5} \\ \frac{27}{10} \end{pmatrix}.$$