

Zur Anwendung von Selektionsindizes bei Zuckerrüben

Von H. GEIDEL und W. HAUFE

(Mit 1 Abbildung)

Zusammenfassung

Ausgehend von den Arbeiten von RAATZ (1894) wurden zunächst für die „klassische“ Theorie der linearen Selektionsindizes praktikable Näherungsformeln abgeleitet und Spezialfälle diskutiert.

Anhand eines Datensatzes (GEIDEL 1982) wurden die abgeleiteten Formeln erläutert und der Zusammenhang mit dem als Selektionskriterium verwendeten „Bereinigten Zuckerertrag (BZE)“ aufgezeigt.

In einem weiteren Abschnitt wurde festgestellt, daß die Raatzsche Wertzahl unter der Annahme, daß die Einzelmerkmale gleiche Bedeutung haben einem Elston-Index entspricht. Für den benutzten Datensatz konnte das bei der Raatzschen Wertzahl implizit verwendete Verhältnis der Heritabilitäten berechnet werden.

Schlüsselworte: Selektionsindex, Zuckerrüben.

On the application of selection indices in sugar beet

Summary

Starting with the papers of RAATZ (1894) to begin with practicable approximation formulae were derived for the “classical” theory of linear selection indices and special cases were discussed.

A data set (GEIDEL 1982) was used to illustrate the derived formulae and the connection with the “Bereinigten Zuckerertrag (BZE)”, which is used as the selection criterion, was shown.

In a further section it was established that the Raatz value corresponds to an Elston-index, under the assumption that the individual characteristics have equal significance. For the data set used it was possible to calculate the ratio of heritabilities which the Raatz value uses implicitly.

Key-words: selection index, sugar beet.

1. Einleitung

Am 1. April 1894 wurde Dr. Wilhelm RAATZ (1864–1919) als erster Saatzucht-leiter bei der Zuckerfabrik Kleinwanzleben angestellt. Die Beschäftigung mit dem Leben und Wirken von W. RAATZ (vgl. HAUFE et al. 1990) führte ganz zwangsläufig dazu, die Anwendung von Selektionsindizes, insbesondere bei Zuckerrüben, einmal systematisch und kritisch zu betrachten.

Ausgangspunkt war die von RAATZ (1894) vorgeschlagene „Wertzahl“, die in der Praxis fast 100 Jahre mit erkennbarem Erfolg benutzt wurde. Eine erste kritische Bewertung dieser Wertzahl erfolgte durch ENDERLEIN (1964).

Die Literatur zu dem Themenkomplex „Selektionsindizes“ ist kaum übersehbar, wenn auch im Zusammenhang mit der Anwendung bei Zuckerrüben nur wenige Veröffentlichungen zu finden sind. Eine umfassende Übersicht hat das AUTORENKOLLEKTIV (1967) gegeben. BAKER (1986) gibt einen Überblick zusammen mit BASIC-Programmen. WRICKE und WEBER (1986) haben gleichfalls die wesentlichen Formeln zusammengestellt. Kaum zu finden sind dagegen Veröffentlichungen über Selektionsexperimente. MANNING (1956) berichtet beispielsweise über Versuche mit Baumwolle.

In der vorliegenden Arbeit sollen nach kurzen allgemeinen Begriffsdefinitionen Anmerkungen zur Anwendung der „klassischen“ Theorie (vgl. SMITH 1936) sowie der Zusammenhang der Raatzschen Wertzahl mit dem Ansatz von ELSTON (1963) aufgezeigt werden.

Dabei muß immer wieder angemerkt werden, daß die Anwendung von Selektionsindizes ein recht allgemeines Problem ist, das nicht nur für die Pflanzen- und Tierzüchtung von Bedeutung ist. Es ist immer relevant, wenn es gilt, aus einer Anzahl von Individuen oder Objekten aufgrund mehrerer festgestellter Merkmale oder Eigenschaften und einer vorgegebenen „Zielvorstellung“ die geeignetsten Individuen oder Objekte auszulesen.

2. Das Selektionsproblem

Allgemein kann von folgender Datenstruktur ausgegangen werden:

Objekt	Merkmal					p
	1	2	.	.	.	
1	x_{11}	x_{12}	.	.	.	x_{1p}
2	x_{21}	x_{22}	.	.	.	x_{2p}
.
.
N	x_{N1}	x_{N2}	.	.	.	x_{Np}

Für N-Objekte oder Individuen liegen für p-Merkmale Werte x_{ij} ($i=1,2, \dots, N$; $j=1, 2, \dots, p$) in Form einer Matrix \mathbf{X} vor.

Beispiele:

- In der Pflanzenzüchtung repräsentiert \mathbf{X} die Matrix phänotypischer Werte.
- Bei einer Weinprämierung enthält die Matrix \mathbf{X} neben chemischen Analysenwerten auch die sensorischen Beurteilungen.
- Bei einer Auswahl einer geeigneten Person für eine bestimmte Aufgabe werden in der Matrix \mathbf{X} beispielsweise neben Schul- bzw. Prüfungsnoten auch persönliche Beurteilungen (z. B. als Rangziffern) zusammengestellt.

Die Zahl der Beispiele läßt sich beliebig vermehren. Deutlich wird dabei aber, daß die Art der Variablen recht unterschiedlich sein kann. Dieses Teilproblem soll hier nur angedeutet, aber nicht vertieft werden.

Das eigentliche Selektionsproblem kann wie folgt beschrieben werden (vgl. u. a. ELSTON 1963):

Vorgegeben sei die oben angegebene Datenstruktur. Wie kann nun eine Rangreihenfolge der N-Individuen/Objekte unter Beachtung aller p-Merkmale/Eigenschaften bei einer vorgegebenen Zielvorstellung erreicht werden?

Eine Lösung dieses Problems ist nur möglich, wenn es gelingt, die Zielvorstellung klar zu definieren.

Das Selektionsproblem ist im Prinzip ein Auswahl- oder auch ein Klassifikationsproblem. Von daher ist es auch verständlich, daß FISHER (1936) die Diskriminanzanalyse als geeignetes Verfahren vorschlug. Diese Anregung griff SMITH (1936) auf und kam so zu einer Lösung, die heute noch als klassisches Verfahren bezeichnet und benutzt wird.

Bezeichnet man die p-Merkmale mit x_i ($i=1, 2, \dots, p$), so kann die Zielvorstellung, d. h. das mögliche Auswahl- oder Beurteilungskriterium, ganz allgemein als eine Funktion Z der x_i dargestellt werden, z. B.

$$Z = f(x_1, x_2, \dots, x_p).$$

Einen solchen Wert Z bezeichnet man als einen Index.

Bei der Funktion $f(x_1, x_2, \dots, x_p)$ sind zunächst zwei Fälle zu unterscheiden:

1. Die Merkmalswerte x_i werden mit zweckmäßig gewählten „Gewichten“ a_i ($i=1, 2, \dots, p$) in einer linearen Funktion zusammengefaßt, d. h. es ist

$$Z = a_1x_1 + a_2x_2 + \dots + a_px_p$$

2. Die Merkmalswerte x_i werden gleichfalls mit zweckmäßig gewählten „Gewichten“ a_i ($i=1, 2, \dots, p$) in einer nichtlinearen Funktion zusammengefaßt, so zum Beispiel

$$Z = \sqrt[n]{x_1^{a_1} x_2^{a_2} \dots x_p^{a_p}}$$

Die „Gewichte“ a_i ($i=1, 2, \dots, p$) sollen die Bedeutung der einzelnen Merkmale für die Zielvorstellung widerspiegeln.

In der Pflanzen- und Tierzucht werden solche Indizes, soweit sie bei der Auslese benutzt werden, als Selektionsindizes bezeichnet.

Da wir in der Pflanzen- und Tierzucht nur die phänotypischen Werte beobachten können, müssen wir überlegen und gegebenenfalls entscheiden, ob wir aufgrund dieser Werte selektieren wollen oder gegebenenfalls weitere Informationen über die nicht direkt beobachtbaren genotypischen Werte mitverwenden.

Betrachten wir hier das übliche lineare Modell für das Merkmal i

$$x_i = g_i + u_i$$

mit

- x_i : phänotypischer Wert,
- g_i : genetisch bedingter Anteil,
- u_i : durch die Umwelt bedingter Anteil,

so erkennen wir, daß Entscheidungen allein aufgrund der phänotypischen Werte x_i zu Verschätzungen führen können. Man wird sich also bemühen, auch Informationen über die nicht direkt beobachtbaren genotypischen Werte zu erhalten und diese dann in einer geeigneten Form zu verwenden.

Damit ergeben sich grundsätzlich vier Fälle für die Konstruktion von Selektionsindizes:

Informationen	Indexfunktion	
	linear	nicht linear
nur phänotypische Schätzwerte	1	2
phänotypische und genotypische Schätzwerte	3	4

Beispiele:

Fallgruppe 1:

- Auswahlkriterien für Personen und Produkte
- DLG-Qualitätskennzahlen
- Bereinigte Polarisation: $BPOL = POL - 0,343 K - 0,343 Na - 0,094 \alpha N - 0,29$
- Bewertungsindex des Bundessortenamtes für Zuckerrübensorten $Z = (RE + POL + BPOL + BZE) / 4$, wobei als Merkmalswerte die Relativzahlen bezogen auf die Standardgruppe benutzt werden.

Fallgruppe 2:

- Bereinigter Zuckerertrag: $BZE = RE * BPOL / 100$
- Index von RAATZ
- ELSTON (1963) hat für diese Gruppe den Fall eines „gewichtsfreien“ Index behandelt.

Fallgruppe 3:

Hier ist insbesondere durch die „klassischen“ Arbeiten von FISHER (1936) und SMITH (1936) die Theorie weitgehend bekannt. Beispiele sind u. a. bei SMITH (1936), MANNING (1956) und WRICKE und WEBER (1986) zu finden.

Fallgruppe 4:

Hier sind nur wenige Ansätze zu finden, z. B. bei ELSTON (1963). Im Prinzip wird versucht, nichtlineare Ansätze durch Transformationen in lineare Funktionen zu überführen und dann die Methoden der Fallgruppe 3 zu benutzen.

3. Die „klassische“ Theorie

Für die „klassische“ Theorie, d. h. die Fallgruppe 3, werden nun unter Berücksichtigung der Arbeit von SMITH (1936) die notwendigen Formeln kurz dargestellt, für einen Spezialfall Formeln für die praktische Anwendung abgeleitet und diese an einem Beispiel erläutert.

3.1 Formeln von SMITH (1936)

Ausgangsbasis ist die oben beschriebene Datenstruktur, wobei für die phänotypischen Werte x_i der p -Merkmale die im vorigen Abschnitt gemachte Annahme eines linearen Modells gelten soll.

Die eigentliche „Zielvorstellung“ ist der genotypische Selektionsindex

$$H = a_1 g_1 + a_2 g_2 + \dots + a_p g_p$$

mit festgelegten, d. h. vorgegebenen, „ökonomischen“ Gewichten a_i . Dies ist die klare Definition des Zuchtziels.

Da die genotypischen Werte g_i nicht direkt bestimmt werden können, muß die eigentliche Auswahl aufgrund eines phänotypischen Selektionsindex

$$I = b_1 x_1 + b_2 x_2 + \dots + b_p x_p$$

erfolgen, wobei die x_i die festgestellten phänotypischen Werte sind.

Die Koeffizienten b_i lassen sich (vgl. SMITH 1936) unter der Annahme, daß die Kovarianz zwischen g_i und u_i Null ist, als Lösungsvektor des folgenden linearen Gleichungssystems

$$\mathbf{Pb} = \mathbf{Ga}$$

d. h.

$$\mathbf{b} = \mathbf{P}^{-1} \mathbf{G} \mathbf{a}$$

ermitteln.

Dabei ist \mathbf{P} die Matrix der phänotypischen Varianzen und Kovarianzen, \mathbf{G} die entsprechende Matrix der genotypischen Varianzen und Kovarianzen und \mathbf{a} der Vektor der vorgegebenen Gewichte der Merkmale.

SMITH (1936) weist darauf hin, daß neben den oben beschriebenen Annahmen (Additivität und Unabhängigkeit der Effekte) vorausgesetzt werden muß, daß H und I normalverteilt sind. Abweichungen von dieser Voraussetzung würden zwar nicht die Berechnung der b_i berühren, wohl aber die bei der Ableitung benutzte Berechnung des genetischen Gewinns durch die Selektion.

3.2 Ableitung von Näherungsformeln für die genetischen Kovarianzen

Da die Bestimmung der Werte der Matrix \mathbf{G} , insbesondere der genetischen Kovarianzen, in der Praxis nicht immer möglich ist, diese Matrix aber für die Bestimmung der b_i erforderlich ist, ist es gegebenenfalls notwendig, mit entsprechenden Näherungswerten zu arbeiten.

Für die Varianzen eines Merkmals gilt unter der bereits oben gemachten Annahme, daß die Kovarianz zwischen g_i und u_i Null ist:

$$\sigma_P^2 = \sigma_G^2 + \sigma_U^2.$$

Daraus errechnet sich die Heritabilität

$$h^2 = \frac{\sigma_G^2}{\sigma_P^2}$$

und

$$e^2 = \frac{\sigma_U^2}{\sigma_P^2} = 1 - h^2.$$

Damit kann man die genetischen Varianzen wie folgt schreiben:

$$\sigma_G^2 = h^2 \sigma_P^2.$$

Eine ähnlich einfache Beziehung ist für die genetischen Kovarianzen nicht bekannt.

Für den phänotypischen Korrelationskoeffizienten zweier Merkmale gilt:

$$r_{Pij} = h_i h_j r_{Gij} + e_i e_j r_{Uij}$$

Aus dieser Beziehung lassen sich unter bestimmten Annahmen Näherungsformeln für die genetischen Kovarianzen ableiten. An dieser Stelle sollen durch drei unterschiedliche Annahmen drei Näherungsformeln für die genetischen Kovarianzen berechnet werden.

a) $r_{Uij} = 0,$

d. h. die Korrelation der Umwelteffekte u_i und u_j sei Null.

Das ergibt

$$G_{ij} = P_{ij}, \quad (a)$$

d. h. die genetischen Kovarianzen sind gleich den phänotypischen Kovarianzen. Diese Annahme (Näherung) hat u. a. auch ELSTON (1963) benutzt.

$$b) r_{G_{ij}} = r_{P_{ij}},$$

d. h. die Korrelation der genotypischen Effekte sei gleich der Korrelation der phänotypischen Werte.

Das ergibt

$$G_{ij} = h_i h_j P_{ij}, \quad (b)$$

d. h. die genotypischen Kovarianzen erhält man aus den phänotypischen Kovarianzen, wenn man diese noch mit $h_i h_j$ multipliziert.

$$c) r_{P_{ij}} = r_{U_{ij}},$$

d. h. die Korrelation der Umwelteffekte sei gleich der Korrelation der phänotypischen Werte.

Das ergibt

$$G_{ij} = (1 - e_i e_j) P_{ij}, \quad (c)$$

d. h. die genotypischen Kovarianzen erhält man aus den phänotypischen Kovarianzen, wenn man diese noch mit $1 - e_i e_j$ multipliziert.

In der Tabelle 1 sind für ausgewählte Werte von h_i und h_j die Multiplikationsfaktoren für die Näherungsformeln (b) und (c) zusammengestellt. Für die Näherungsformel (a) ist dieser Faktor immer 1.

Für $h_i = h_j$ (Diagonale in Tab. 1) stimmen die beiden Näherungsformeln (b) und (c) überein. Allgemein gilt

$$1 - e_i e_i \geq h_i h_j.$$

Tabelle 1

Multiplikationsfaktoren für die Näherungsformel (b) oberhalb der Diagonale und (c) unterhalb der Diagonale

h_i h_j	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
0	0	0	0	0	0	0	0	0	0	0	0
0,1	0,01	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,10
0,2	0,02	0,03	0,04	0,06	0,08	0,10	0,12	0,14	0,16	0,18	0,20
0,3	0,05	0,05	0,07	0,09	0,12	0,15	0,18	0,21	0,24	0,27	0,30
0,4	0,08	0,09	0,10	0,13	0,16	0,20	0,24	0,28	0,32	0,36	0,40
0,5	0,13	0,14	0,15	0,17	0,21	0,25	0,30	0,35	0,40	0,45	0,50
0,6	0,20	0,20	0,22	0,24	0,27	0,31	0,36	0,42	0,48	0,54	0,60
0,7	0,29	0,29	0,30	0,32	0,35	0,38	0,43	0,49	0,56	0,63	0,70
0,8	0,40	0,40	0,41	0,43	0,45	0,48	0,52	0,57	0,64	0,72	0,80
0,9	0,56	0,57	0,57	0,58	0,60	0,62	0,65	0,69	0,74	0,81	0,90
1,0	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

3.3 Formeln unter der Annahme $G_{ij} = h_i h_j P_{ij}$

Im folgenden soll die Annahme (b), d. h.

$$G_{ij} = h_i h_j P_{ij}$$

als Näherungsformel für die genotypischen Kovarianzen benutzt werden. Damit ist

$$\mathbf{G} = \begin{pmatrix} h_1^2 P_{11} & h_1 h_2 P_{12} & \dots & h_1 h_p P_{1p} \\ h_1 h_2 P_{12} & h_2^2 P_{22} & \dots & h_2 h_p P_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ h_1 h_p P_{1p} & h_2 h_p P_{2p} & \dots & h_p^2 P_{pp} \end{pmatrix}$$

Mit einer Matrix \mathbf{H}

$$\mathbf{H} = \begin{pmatrix} h_1 & 0 & \dots & 0 \\ 0 & h_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h_p \end{pmatrix}$$

läßt sich \mathbf{G} wie folgt schreiben:

$$\mathbf{G} = \mathbf{H} \mathbf{P} \mathbf{H}$$

und damit \mathbf{b}

$$\mathbf{b} = \mathbf{P}^{-1} \mathbf{H} \mathbf{P} \mathbf{H} \mathbf{a}.$$

Aus dieser Beziehung lassen sich nun für wesentliche Spezialfälle relativ einfache Formeln ableiten.

a) $h_1 \approx h_2 \approx \dots \approx h_p \approx h$

Unter der Annahme, daß die Heritabilitäten als gleich vorausgesetzt werden können, gilt

$$\mathbf{b} = h^2 \mathbf{a},$$

d. h. die gesuchten Koeffizienten b_i ($i = 1, 2, \dots, p$) entsprechen bis auf den Faktor h^2 den angenommenen Gewichten a_i ($i = 1, 2, \dots, p$).

b) $P_{ij} = 0$

Unter der Annahme, daß die Kovarianzen Null sind oder praktisch vernachlässigt werden können, gilt

$$\mathbf{b} = \mathbf{H}^2 \mathbf{a},$$

d. h. die angenommenen Gewichte a_i ($i = 1, 2, \dots, p$) sind mit den entsprechenden Heritabilitäten h_i^2 ($i = 1, 2, \dots, p$) zu multiplizieren.

3.4 Formeln für den Fall $p = 2$

Da in vielen Fällen Selektionsindizes nur für zwei wesentliche Merkmale benötigt werden, werden in diesem Abschnitt die entsprechenden Formeln zusammengestellt. Dabei lassen sich die in den Abschnitten 3.2 und 3.3 angegebenen Matrizen und Vektoren mit erträglichem Aufwand direkt berechnen.

Es ist

$$\mathbf{P} = \begin{pmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{pmatrix}$$

und damit

$$\mathbf{P}^{-1} = \frac{1}{P_{12}^2 - P_{11}P_{22}} \begin{pmatrix} -P_{22} & P_{12} \\ P_{12} & -P_{11} \end{pmatrix}$$

Für \mathbf{G} gilt mit der Näherung (b) für die Kovarianzen

$$\mathbf{G} = \begin{pmatrix} h_1^2 P_{11} & h_1 h_2 P_{12} \\ h_1 h_2 P_{12} & h_2^2 P_{22} \end{pmatrix}$$

Mit den Gewichten a_1 und a_2 ist der genotypische Selektionsindex

$$H = a_1 g_1 + a_2 g_2$$

und entsprechend mit b_1 und b_2 der phänotypische Selektionsindex

$$I = b_1 x_1 + b_2 x_2.$$

Für den Vektor \mathbf{b} der gesuchten Koeffizienten b_1 und b_2 gilt dann

$$\mathbf{b} = \frac{1}{P_{12}^2 - P_{11}P_{22}} \begin{pmatrix} h_1(h_2 P_{12}^2 - h_1 P_{11}P_{22}) a_1 + h_2 P_{22} P_{12}(h_2 - h_1) a_2 \\ h_1 P_{11} P_{12}(h_1 - h_2) a_1 + h_2(h_1 P_{12}^2 - h_2 P_{11}P_{22}) a_2 \end{pmatrix}$$

Da es bei den Gewichten a_1 und a_2 nur auf das Verhältnis ankommt, kann man folgende Vereinfachung vornehmen:

$$a = \frac{a_2}{a_1}$$

d. h. es ist $a_2 = a a_1$, soweit $a_1 \neq 0$ ist.

Auch bei den beiden Heritabilitäten h_1 und h_2 ist es sinnvoll, mit dem Verhältnis

$$h = \frac{h_2}{h_1}$$

zu arbeiten, soweit $h_1 \neq 0$ ist. D. h. es ist $h_2 = h h_1$.

Da auch bei den Koeffizienten b_1 und b_2 letztlich nur das Verhältnis dieser beiden Größen interessiert, berechnet man, soweit $b_1 \neq 0$ ist, zweckmäßigerweise einen Vektor \mathbf{b}^* , d. h.

$$\mathbf{b}^* = \begin{pmatrix} 1 \\ \frac{b_2}{b_1} \end{pmatrix} = \begin{pmatrix} b_1^* \\ b_2^* \end{pmatrix}$$

und erhält so schließlich den wesentlichen Koeffizienten b_2^* :

$$b_2^* = \frac{P_{11} P_{12} (1 - h) + h (P_{12}^2 - h P_{11} P_{22}) a}{h P_{12}^2 - P_{11} P_{22} + h P_{22} P_{12} (h - 1) a}.$$

b_2^* ist bei einer gegebenen Matrix \mathbf{P} eine Funktion von h und a , d. h. $b_2^* = f(h, a)$.

Für spezielle Werte von h kann man unter den gemachten Annahmen die Werte b_2^* wie folgt angeben:

$$\text{a) } h = 0$$

$$b_2^* = -\frac{P_{12}}{P_{22}}$$

b) $h=1$
 $b_2^* = a$

c) $h=\infty$

$$b_2^* = -\frac{P_{11}}{P_{12}}$$

Die Beziehung für b_2^* kann u. a. benutzt werden, um mögliche Veränderungen der b_2^* -Werte und damit auch Veränderungen des phänotypischen Selektionsindex I in Abhängigkeit von möglichen Schwankungen (bzw. Veränderungen) von \mathbf{P} , h und a zu untersuchen.

Die in Abschnitt 3.3 behandelten Spezialfälle führen auch für $p=2$ zu relativ einfachen Formeln.

3.5 Beispiel: Bereinigter Zuckerertrag

Für den bereinigten Zuckerertrag (BZE) gilt:

$$\text{BZE} = \text{RE} \cdot \text{BPOL} / 100$$

mit RE: Rübenertrag und BPOL: bereinigte Polarisation.

Dies ist (vgl. Abschnitt 2, Fallgruppe 2) ein nichtlinearer Selektionsindex auf der Basis nur von phänotypischen Werten. In diesem Fall ist es aber recht einfach, durch Logarithmieren eine lineare Funktion zu erhalten, für die die in Abschnitt 3.4 abgeleiteten Formeln benutzt werden können.

Es ist

$$\lg \text{BZE} = \lg \text{RE} + \lg \text{BPOL} - \lg 100$$

Mit den Variablen

$$\begin{aligned} x_1 &= \lg \text{RE} \\ x_2 &= \lg \text{BPOL} \\ x_3 &= \lg \text{BZE} \end{aligned}$$

erhält man

$$x_3 = x_1 + x_2 - 2.$$

Die Konstante $\lg 100=2$ kann an dieser Stelle zunächst unberücksichtigt bleiben, da sie keinen Einfluß auf die Rangreihenfolge hat.

Für das Zuchtziel „Bereinigter Zuckerertrag“ ist somit

$$x_3 = x_1 + x_2.$$

Durch diese Beziehung sind die Gewichte $a_1 = a_2 = 1$ und damit auch $a = 1$ festgelegt.

Für den phänotypischen Selektionsindex $I = x_1 + b_2^* \cdot x_2$ ist der Koeffizient b_2^* dann:

$$b_2^* = \frac{P_{11} P_{12} + h(P_{12}^2 - P_{11} P_{12}) - h^2 P_{11} P_{22}}{-P_{11} P_{22} + h(P_{12}^2 - P_{22} P_{12}) + h^2 P_{22} P_{12}},$$

d. h. eine Funktion der Elemente der Matrix \mathbf{P} , der phänotypischen Varianzen und Kovarianzen, und von h , dem Verhältnis der beiden Heritabilitäten h_1 und h_2 .

Als Datensatz (vgl. GEIDEL 1982) wird eine Zuckerrübenpopulation ($n=69$ Einzelrüben) benutzt. Es ist dies eine typische „Stutzpopulation“ einer sogenannten Stammzuchtparzelle, bei der die extremen Werte (Fehlstellennachbarn, doppelt besetzte Pflanzstellen, kranke Pflanzen, extrem abweichende Typen, extrem kleine Rüben) bereits eliminiert wurden. Dabei muß festgehalten werden, daß dieser Datensatz hier nur die Rechengänge erläutern soll. Gegebenenfalls wäre es nützlich, durch die Verarbeitung umfangreicherer Daten bzw. durch entsprechende Simulationen mögliche Änderungen bei b_2^* zu untersuchen.

In Tabelle 2 sind für die 20 besten Objekte aufgrund des bereinigten Zuckerertrages (BZE) die Werte Rübenenertrag (RE), bereinigte Polarisierung (BPOL) und der bereinigte Zuckerertrag (BZE) sowie die Mittelwerte und Standardabweichungen aus dem Gesamtmaterial ($n=69$) zusammengestellt.

Tabelle 2
Daten von 20 Objekten (vgl. GEIDEL 1982)

Objekt	RE [g]	BPOL [%]	BZE [g]
1	1220	13,34	162,8
2	1115	13,88	154,8
3	1115	13,28	148,1
4	935	15,61	146,0
5	1195	11,92	142,4
6	980	12,82	125,6
7	865	14,45	125,0
8	850	14,04	119,3
9	860	13,83	118,9
10	855	13,72	117,3
11	725	15,41	111,7
12	730	14,45	105,5
13	780	13,44	104,8
14	730	13,08	95,5
15	715	13,26	94,8
16	730	12,80	93,4
17	665	14,04	93,3
18	705	12,87	90,7
19	680	13,17	89,6
20	640	13,84	88,6
\bar{x}	595	12,84	77,5
s	217	1,12	31,4

RE: Rübenenertrag [g]
BPOL: bereinigte Polarisierung [%]
BZE: bereinigter Zuckerertrag [g]

Für die logarithmierten Werte $x_1 = \lg \text{RE}$ und $x_2 = \lg \text{BPOL}$ des Datensatzes erhält man folgende Matrix \mathbf{P} :

$$\mathbf{P} = \frac{1}{1000} \begin{pmatrix} 20,07 & 0,533 \\ 0,533 & 0,512 \end{pmatrix}$$

Dabei ist bemerkenswert, aber für sogenannte „Stutzpopulationen“ nicht überraschend, daß die Kovarianz positiv ist.

Für die Heritabilitäten werden folgende Werte angenommen:

$$h_1 = 0,5$$

$$h_2 = 0,75$$

d. h. es ist $h = 1,5$.

Mit diesen Werten erhält man

$$b_2^* = 2,91.$$

In der Tabelle 3 sind für die Daten aus Tabelle 2 die Werte für $x_1 = \lg RE$, $x_2 = \lg BPOL$, der jeweils berechnete Wert des Selektionsindex $I = x_1 + 2,91 x_2$ sowie die aufgrund von I ermittelten Rangzahlen zusammengestellt.

Tabelle 3

Transformierte Werte, Indexwerte und Rangzahlen für die Daten aus Tabelle 2

Objekt	$x_1 = \lg RE$	$x_2 = \lg BPOL$	$I = x_1 + 2,91 x_2$	Rang
1	3,0864	1,1252	6,3607	3
2	3,0473	1,1424	6,3717	2
3	3,0473	1,1232	6,3158	5
4	2,9708	1,1934	6,4436	1
5	3,0774	1,0763	6,2094	12
6	2,9912	1,1079	6,2152	11
7	2,9370	1,1599	6,3123	6
8	2,9294	1,1474	6,2683	7
9	2,9345	1,1408	6,2542	8
10	2,9320	1,1374	6,2418	9
11	2,8603	1,1878	6,3168	4
12	2,8633	1,1599	6,2386	10
13	2,8921	1,1284	6,1757	13
14	2,8633	1,1166	6,1126	17
15	2,8543	1,1225	6,1208	16
16	2,8633	1,1072	6,0853	19
17	2,8228	1,1474	6,1617	14
18	2,8482	1,1096	6,0771	20
19	2,8325	1,1196	6,0905	18
20	2,8062	1,1411	6,1268	15

Man erkennt, daß die Rangzahlen aufgrund des Selektionsindex I sich zum Teil recht beachtlich von den Rangzahlen aufgrund des bereinigten Zuckerertrages (Reihenfolge der Objekte, vgl. Tab. 2) unterscheiden.

Da auch der „Bereinigte Zuckerertrag (BZE)“ ein Selektionsindex ist (vgl. Abschnitt 2, Fallgruppe 2), soll an dieser Stelle zunächst untersucht werden, unter welchen Bedingungen die beiden Selektionsindizes, d. h. I und BZE, übereinstimmen.

Dazu kann man u. a. von folgenden Gleichungen ausgehen:

$$I = \lg RE + b_2^* \lg BPOL \quad (-2)$$

$$\lg BZE = \lg RE + \lg BPOL - 2$$

Diese beiden Gleichungen stimmen überein, d. h. sind gleich, wenn $b_2^* = 1$ ist.

Aus $b_2^* = 1$ folgt die nachstehende quadratische Gleichung für h :

$$P_{11} P_{12} + h (P_{12}^2 - P_{11} P_{12}) - h^2 P_{11} P_{22} = -P_{11} P_{22} + h (P_{12}^2 - P_{22} P_{12}) + h^2 P_{22} P_{12}$$

bzw.

$$h^2 + h \frac{P_{12}(P_{11} - P_{22})}{P_{22}(P_{11} + P_{12})} - \frac{P_{11}(P_{22} + P_{12})}{P_{22}(P_{11} + P_{12})} = 0$$

Die Nullstellen dieser quadratischen Gleichung lassen sich ganz allgemein angeben. Es sind

$$h^* = 1$$

und

$$h^{**} = - \frac{P_{11}(P_{22} + P_{12})}{P_{22}(P_{11} + P_{12})}$$

Mit den obigen Werten der Matrix \mathbf{P} erhält man:

$$h^* = 1 \text{ und } h^{**} = -1,99.$$

Diese Ergebnisse sind wie folgt zu interpretieren:

- a) ein negatives Verhältnis der beiden Heritabilitäten ist nicht sinnvoll.
- b) $h^* = 1$ bedeutet: $h_1 = h_2$, d. h. die beiden Heritabilitäten werden bei Verwendung von BZE als gleich angenommen.

Aus der obigen Beziehung für h^{**} kann allgemein abgeleitet werden, unter welchen Voraussetzungen $h^{**} > 0$ werden kann, d. h. gegebenenfalls eine zweite positive Nullstelle zu betrachten wäre.

Bei $P_{12} < 0$ wird $h^{**} > 0$ wenn

- a) $|P_{12}| > P_{11}$ und $|P_{12}| < P_{22}$
- b) $|P_{12}| < P_{11}$ und $|P_{12}| > P_{22}$

Da die phänotypischen Varianzen und Kovarianzen und erst recht die entsprechenden genotypischen Varianzen und Kovarianzen sowie die abgeleiteten Heritabilitäten „fehlerbehaftet“ sind, empfiehlt es sich, die Stabilität der Koeffizienten b_i bzw. b_i^* zu untersuchen.

Im vorliegenden Fall, d. h. bei $a=1$, ist b_2^* bei vorgegebenen Werten der Matrix \mathbf{P} nur noch eine Funktion von h .

In der Tabelle 4 sind einige Werte der Funktion $b_2^* = f(h)$ bei der vorgegebenen Matrix \mathbf{P} zusammengestellt.

Tabelle 4

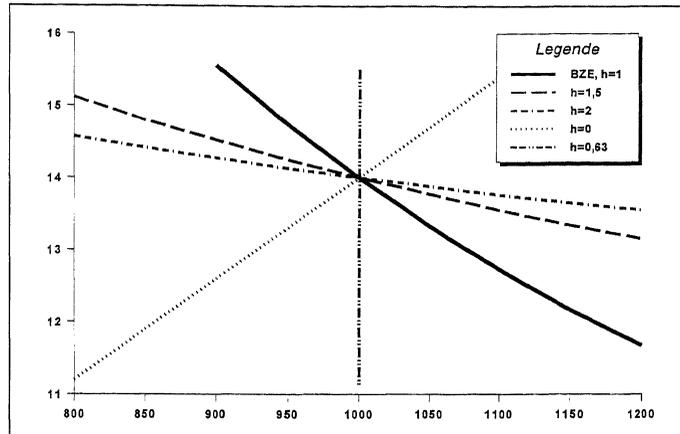
Werte der Funktion $b_2^* = f(h)$

h	b_2^*	h	b_2^*	h	b_2^*
0	-1,04	1,5	2,91	3,5	22,00
0,5	-0,29	2,0	5,59	4,0	33,31
0,63	0	2,5	9,31	6,12	∞
1,0	1	3,0	14,52	∞	-37,65

An dieser Stelle muß noch einmal festgestellt werden, daß b_2^* nicht direkt von h_1 und h_2 , sondern nur von dem Verhältnis dieser beiden Größen abhängt.

Um den Einfluß von h auf die Selektionsfunktion deutlich zu machen, wurden für einige Werte der Tabelle 4 die entsprechenden Funktionen in ein RE,

Abb. 1: Selektionsfunktionen in Abhängigkeit von h
(Ordinate BPOL [%],
Abszisse RE [g])



BPOL-Diagramm (vgl. Abb. 1 eingezeichnet. Die Funktionen wurden so normiert, daß sie alle durch den Punkt RE=1000 und BPOL=14,0 gehen.

Aus dem Diagramm lassen sich einige Schlußfolgerungen, die natürlich nur für den hier betrachteten Datensatz Gültigkeit haben, ableiten:

- a) $h=0,63$, d. h. $b_2^*=0$
Eine Selektion erfolgt nur aufgrund des Ertrages
- b) $h=1$, d. h. $b_2^*=1$
Selektionsindex I und BZE stimmen überein
- c) $h > 1$

Die bereinigte Polarisation wird stärker berücksichtigt als bei einer Selektion aufgrund des bereinigten Zuckerertrages (BZE).

4. Vergleich der Raatzschen Wertzahl mit dem Ansatz von ELSTON

4.1 Der Ansatz von ELSTON

Da es in der Praxis unter Umständen schwierig ist, Schätzwerte für die genetischen Varianzen und Kovarianzen zu ermitteln, ist es verständlich, daß nach anderen Lösungsmöglichkeiten für die Konstruktion von Selektionsindizes gesucht wird. So hat ELSTON (1963) einen Ansatz veröffentlicht, der von einer nichtlinearen Selektionsfunktion ausgeht. Dabei bezieht er sich auf das ökonomische Prinzip der Substitution. Sein Index kann wie folgt formuliert werden:

$$I = \prod_{i=1}^p (x_i - k_i)$$

Die k_i sind dabei untere Grenzen der einzelnen Merkmale, die nicht unterschritten werden sollen.

Es zeigt sich, daß dieser Index im Prinzip „gewichtsfrei“ angewandt werden kann. Es wird angenommen, daß die einzelnen Merkmale etwa gleich bedeutsam sind. Dies kann gegebenenfalls durch Standardisierung und Transformierung erreicht werden.

4.2 Die Raatzsche Wertzahl

RAATZ (1894) hatte folgende Wertzahl als Selektionsindex vorgeschlagen, der für die Selektion von Individuen (Mutterrüben) vorgesehen war:

$$I = 10\sqrt{z_1 \cdot z_2}$$

Dabei waren z_1 und z_2 wie folgt festgelegt:

$$z_1 = 10 + \frac{x_1 - \bar{x}_1}{s_1} \cdot p \quad (\text{Rübenertrag})$$

$$z_2 = 10 + \frac{x_2 - \bar{x}_2}{s_2} \cdot q \quad (\text{Zuckergehalt})$$

Mit dem Verhältnis: $p:q$ drückte RAATZ die unterschiedliche Bedeutung der beiden Merkmale für die verschiedenen Zuchtrichtungen aus.

Die Wertzahl war so normiert, daß den Mittelwerten \bar{x}_1 und \bar{x}_2 ein Indexwert von 100 entsprach. Diese Normierung ermöglichte es, daß die Auswertung mit Hilfe von sogenannten Wertzahlentabellen (vgl. HAUFE et al. 1990) erfolgen konnte.

Die von RAATZ vorgenommene Standardisierung der Meßwerte erfolgte mit dem gleichen Ziel, das auch ELSTON anstrebte, daß nämlich die einzelnen Merkmale etwa gleich bedeutsam sein sollen.

Wenn man davon ausgeht, daß für z_1 und z_2 die unteren Grenzen k_1 und k_2 jeweils Null sind, so entspricht die Raatzsche Wertzahl einem Elston-Index. Die Wurzelziehung und die Multiplikation mit 10 ändern nicht die Rangreihenfolge gegenüber einem Index aus dem einfachen Produkt $z_1 z_2$.

Tabelle 5

$z_1 z_2$, die Raatzsche Wertzahl und Rangzahlen für die Daten aus Tabelle 2

Objekt	z_1	z_2	$I = 10\sqrt{z_1 z_2}$	Rang
1	12,88	10,45	116,0	3
2	12,40	10,93	116,4	2
3	12,40	10,39	113,5	5
4	11,57	12,47	120,1	1
5	12,76	9,18	108,2	12
6	11,77	9,98	108,4	11
7	11,24	11,44	113,4	6
8	11,18	11,07	111,2	7
9	11,22	10,88	110,5	8
10	11,20	10,79	109,9	10
11	10,60	12,29	114,1	4
12	10,62	11,44	110,2	9
13	10,85	10,54	106,9	13
14	10,62	10,21	104,1	17
15	10,55	10,38	104,6	16
16	10,62	9,96	102,8	19
17	10,32	11,07	106,9	14
18	10,51	10,03	102,7	20
19	10,39	10,29	103,4	18
20	10,21	10,89	105,4	15

4.3 Beispiel: Bereinigter Zuckerertrag

In der Tabelle 5 sind für die Daten der Tabelle 2 die Werte z_1 und z_2 , die Raatzsche Wertzahl I sowie die sich aufgrund von I ergebenden Rangzahlen zusammengestellt. Dabei wurde $p=q=1$ gesetzt, d. h. die Merkmale wurden als gleichwertig betrachtet.

Man erkennt, daß die Rangzahlen aufgrund der Raatzschen Wertzahl sich ebenfalls recht beachtlich von den Rangzahlen aufgrund des bereinigten Zuckerertrages (Reihenfolge der Objekte vgl. Tab. 2) unterscheiden. Überraschend ist aber eine gute Übereinstimmung mit den Rangzahlen der Tabelle 3.

Zumindest für den vorliegenden Datensatz können die Raatzsche Wertzahl und der in Abschnitt 3 abgeleitete Index als gleich bewertet werden.

Aus einem Vergleich der entsprechenden Selektionsfunktionen läßt sich für die Raatzsche Wertzahl für das Verhältnis der Heritabilitäten h_1 und h_2 im Mittel ein Wert von

$$h=1,54$$

berechnen. Die Selektionsfunktion aufgrund der Raatzschen Wertzahl unterscheidet sich daher nur unwesentlich von der Selektionsfunktion für $h=1,5$ in Abbildung 1.

Literatur

- AUTORENKOLLEKTIV, 1967: Gleichzeitige Selektion nach mehreren Merkmalen und deren Anwendung in der Pflanzen- und Tierzüchtung. Deutsche Akademie der Landwirtschaftswissenschaften, Berlin.
- BAKER, R. J., 1986: Selection indices in plant breeding. CRS Press, Inc. Boca Raton, Florida.
- ELSTON, R. C., 1963: A weight-free index for the purpose of ranking or selection with respect to several traits at a time. *Biometrics* 19, 85–97.
- ENDERLEIN, G., 1964: Die Bedeutung von Wertindizes für die Selektion. *Biom. Zeitschr.* 6, 217–245.
- FISHER, R. A., 1936: The use of multiple measurements in taxonomic problems. *Ann. of Eugenics* 7, 179–189.
- GEIDEL, H., 1982: Zur Anwendung von Selektionsindizes, insbesondere bei Zuckerrüben. Bericht über die Arbeitstagung 1982 der „Arbeitsgemeinschaft der Saatzuchtleiter“ in Gumpenstein, 175–189.
- HAUFE, W., H. GEIDEL und E. JUNGHANS, 1990: Wilhelm Raatz 1894–1919. KWS Kleinwanzlebener Saatzucht AG, Einbeck.
- MANNING, H. L., 1956: Yield improvement from a selection index technique with cotton. *Heredity* 10, 303–322.
- RAATZ, W., 1894: Mathematische Vorstudien für ein Punktiertssystem. In: Botanische Berichte. Archiv Kleinwanzleben.
- SMITH, H. F., 1936: A discriminant function for plant selection. *Ann. of Eugenics* 7, 240–250.
- WRICKE, G. and W. E. WEBER, 1986: Quantitative Genetics and Selection in Plant Breeding. Berlin.

(Manuskript eingelangt am 2. Mai 1994, angenommen am 16. August 1994)

Anschrift der Verfasser:

Prof. Dr. Hans GEIDEL, Birkheckenstraße 100 A, D-70599 Stuttgart und Werner HAUFE, An der Allee 9, D-32139 Spenge