Diskriminanzanalyse

mit ausformulierten Herleitungen

Hans Stocker, Schaffhausen
unter Mithilfe von Kurt Hug, Le Landeron

Version 16.01.2023

Contents

1 Einführung

Wer sich mit Diskriminanzanalyse beschäftigt stösst immer wieder auf verschiedene Wege zur Herleitung der Lösung der Diskriminanzfunktion. Diese Beweisführungen und Herleitungen werden in diesem Aufsatz ausführlich zusammengefasst und beschrieben. Kapitel 2 beschreibt die Ziele und die Notation der linearen Diskriminanzanalyse (LDA). Der Bezug zur Varianzanalyse und die Formeln zur Bestimmung der Diskriminanzkoeffizienten mittels Lösung des Rayleigh-Quotienten nach dem Satz von Courant-Fischer werden hergeleitet und sind in den Appendices A und B ausführlich beschrieben.

keywords

LDA, linear discriminant analysis, ausformulierte Herleitungen, formulated derivations, proof with Courant-Fischer theorem.

2 Lineare Diskriminanzanalyse LDA

2.1 Einführung in die LDA

Nach Kosfeld, R (2018), Skript Vorlesung 'Multivariate Statistik', Universität Kassel [1].

Die lineare **Diskriminanzanalyse**, **LDA**, ist eine statistische Methode, welche einerseits die **Diskrimination** und andrerseits die **Klassifikation** von **Gruppen von Daten (mit verschiedenen Objekten)** ermöglicht und zwei Ziele verfolgt:

1. Angewandt zur **Diskrimination** *(zum Trennen)* können mittels LDA die beiden folgenden Fragestellungen untersucht werden:

- Können die vorgegebenen Gruppen mit Hilfe metrisch skalierter Variablen **signifikant** voneinander getrennt bezw. erklärt werden? D.h. unterscheiden sich die Gruppen signifikant?
- Welche der metrisch skalierten Variablen sind für die Trennung der Gruppen von besonderer Relevanz? D.h. wie lassen sich die Gruppenunterschiede erklären?

2. Angewandt zur **Klassifikation** *(zum Einordnen)* erhält man mittels LDA Antworten auf die folgende Frage:

- In welcher Gruppe ist eine neue Untersuchungseinheit (d.h. ein neuer Beobachtunswert), deren Gruppenzugehörigkeit nicht bekannt ist, aufgrund ihrer Merkmalsausprägungen bei den metrisch skalierten Merkmalsvariablen einzuordnen.

Die Idee der LDA *(genauer der linearen multivariaten Diskriminanzanalyse)* besteht darin, mehrere Merkmalsvariablen unter Wahrung eines minimalen Informationsverlustes durch eine Linearkombination einer einzigen Variable zu erfassen.

Notation der Variablen und Indizes

Grösse	*Abk.*	*Index*	*max.Wert*	*Bereich*
Gruppe	G	j	g	$1 \leq j \leq g$
Objekt	O	i	n_j	$1 \leq i \leq n_j$
Merkmal	X	k	m	$1 \leq k \leq m$

Allgemeiner Fall

Gegeben seien Datensätze von Merkmalsvariablen verschiedener Gruppen. Die Merkmalsvariablen, die an einzelnen Objekten gemessen werden, sind z.B. gemäss einer oft zitierten Anwendung Längen und Breiten der Kron- und Kelchblätter verschiedener Iris-Arten (Gruppen): sepal length, sepal width, petal length, petal width. Die untersuchten Iris-Arten (Gruppen) seien I. setosa, I. versicolor, I. virginica. Eine Linearkombination der Merkmalsvariablen, die diese Gruppen optimal trennt heisst **Diskriminanzfunktion** (Trennfunktion).

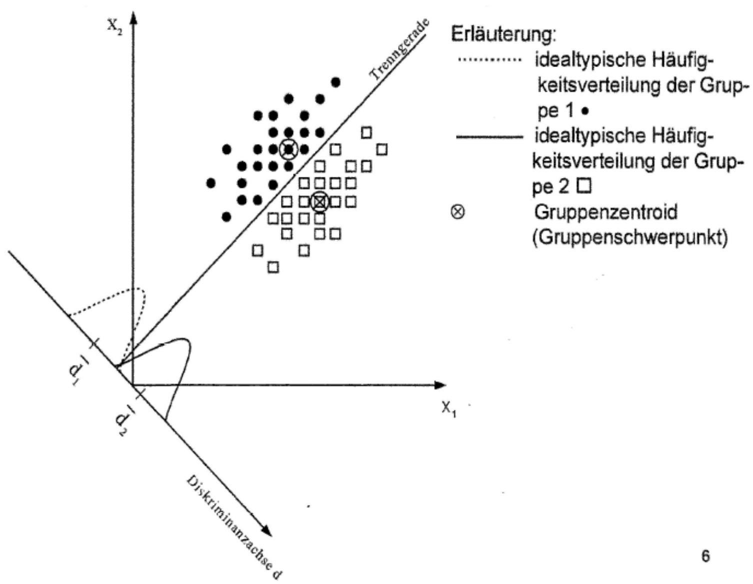
Im allgemeinen Fall lautet diese:

$$D = a_0 + a_1X_1 + a_2X_2 + ... + a_jX_k \tag{2-1}$$

wobei gilt:

D = Diskriminanzfunktion
X_k = Merkmalsvariable, kurz Merkmal mit $(k = 1, ..., m)$
a_j = Diskriminanzkoeffizient für Merkmalsvariable X_k
a_0 = Konstantes Glied

Abb. 1: Geometrische Idee der Diskriminanzanalyse:



Die Linearkombination, die die beiden Merkmale X_1 und X_2 optimal trennen soll, lautet im Zwei-Gruppen-Fall:

$$D = a_0 + a_1X_1 + a_2X_2 \tag{2-2}$$

Geometrisch interpretiert, wird das Koordinatensystem um den Winkel α gedreht, wobei α den Winkel zwischen der alten und der neuen Achse darstellt (siehe Abb. 1). Gemäss der Formel der Rotation des Koordinatensystems ist $a_1 = \cos\alpha$ und $a_2 = \sin\alpha$. Es gilt $a_1^2 + a_2^2 = 1$.

2.2 Berechnung der Diskriminanzwerte d_i

Die Funktion D definiert die Diskriminanzachse d (siehe Abb. 1). Die orthogonale Projektion der individuellen Beobachtungswerte der Objekte auf d ergibt die

Diskriminanzwerte d_i (engl. scores) der Objekte auf dieser Achse. Diese sind gegeben durch:

$$d_i = a_0 + a_1 x_{i1} + ... + a_2 x_{ik} \tag{2-3}$$

mit der Verschiebung aus dem Ursprung a_0 und den Beobachtungswerten $x_{i1}, ..., x_{ik}$ des i-ten Objektes. Diese Schreibweise bietet sich für die Untersuchung mit noch nicht klassifizierten Objekten an.

Allgemein gilt dann:

x_{jik} für den Beobachtungswertdes i-ten Objektes in der j-ten Gruppe beim Merkmal X_k, also die Werte auf den ursprünglichen Achsen,

sowie

d_{ji} für den Diskriminanzwert des i-ten Objektes in der j-tenGruppe
mit $j = 1, ..., g$ Gruppen
und $i = 1, ..., n_j$ Objekten pro Gruppe

ergibt die Diskriminanzfunktion für den Fall mit zwei Merkmalen:

$$d_{ji} = a_0 + a_1 x_{ji1} + a_2 x_{ji2} \tag{2-4}$$

Beispiel: Im berühmten Iris Datensatz von Fisher können z.B. sinnvollerweise die Menge {Blütenlänge, Blütenbreite, Kelchlänge, Kelchbreite} den Merkmalsvariablen (engl. features) und die Menge der Arten {setosa, versicolor, virginica} den Gruppen (engl. classes) zugeordnet werden.

z.B. $x_{jik} = x_{(class\ setosa,\ sample\ \#20,\ sepal\ length)}$

Nomenklatur (auf Diskriminanzachse)

Ziel der LDA ist es, die m Merkmale $X_k (k = 1, ..., m)$ der g Gruppen $G_j (j = 1, ..., g)$ mit je n_g Objekten $O_i (i = 1, ..., n_g)$ zu reduzieren und durch die eindimensionale Grösse d_{ji} auszudrücken, also

Gruppe 1: $d_{11},\ d_{12}, ..., d_{1n_1}$ mit Mittelwert $\bar{d_1}$
Gruppe 2: $d_{21},\ d_{22}, ..., d_{2n_2}$ mit Mittelwert $\bar{d_2}$
$\vdots$
Gruppe j: $d_{j1},\ d_{j2}, ..., d_{jn_j}$ mit Mittelwert $\bar{d_j}$
$\vdots$
Gruppe g: $d_{g1},\ d_{g2}, ..., d_{gn_g}$ mit Mittelwert $\bar{d_g}$.

Bei Beschränkung auf Zwei-Gruppen-Fall:

Gruppe 1: $d_{11},\ d_{12}, ..., d_{1n_1}$ mit Mittelwert $\bar{d_1}$
Gruppe 2: $d_{21},\ d_{22}, ..., d_{2n_2}$ mit Mittelwert $\bar{d_2}$.

2.3 Lösung der Diskriminanzfunktion

2.3.1 Prinzip gemäss ANOVA

LDA und ANOVA (Analysis of Variance) sind bezüglich der mathematischen Prinzipien identisch. Das Prinzip der Varianzanalyse wird in Appendix A näher erläutert und lässt sich sinngemäss direkt auf die LDA übertragen, wovon wir hier Gebrauch machen werden.

Appendix A

Mit der einfache Varianzanalyse (ANOVA = Analysis of Variance) wird die Hypothese geprüft, ob die Mittelwerte zweier oder mehrerer Stichproben identisch sind, welche aus normalverteilten Grundgesamtheiten gezogen werden und welche denselben Mittelwert besitzen.

Für eine Variable X mit mehreren Gruppen g gilt allgemein:

$$\sum_{j=1}^{g}\sum_{i=1}^{n_j}(x_{ji}-\bar{x})^2 = \sum_{j=1}^{g} n_j(\bar{x}_j-\bar{x})^2 + \sum_{j=1}^{g}\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)^2 \tag{2-5}$$

Dies entspricht der Streuungszerlegung

$$T_S = B_S + W_S, \tag{2-6}$$

mit
T_S = Gesamtvarianz
B_S = Varianz (Streung) zwischen (between) den Mittelwerten der Gruppen
W_S = gepoolte* Varianz(Streung) innerhalb (within) der Gruppen

* gepoolt bedeutet hier: Varianz der Gruppen zusammengeführt

Entsprechend lautet diese Zerlegung in der LDA:

$$T_d = B_d + W_d. \tag{2-7}$$

In Erweiterung des bekannten t-Tests wird das Verhältnis der Varianz zwischen den Mittelwerten der Gruppen und der gepoolten Varianzen innerhalb der Gruppen mit einem F-Test auf Signifikanz geprüft. Dieses bei der ANOVA verwendete Verhältnis wird bei der LDA als Diskriminanzkriterium = λ bezeichnet. Bei guter Trennung ist die Varianz zwischen den Mittelwerten möglichst gross und die Varianz innerhalb der Gruppen möglichst klein $\Rightarrow$ gesucht ist das Maximum von λ. Anstelle der Werte x der Variablen X werden für die LDA die Werte d der Diskriminanzfunktion D entsprechend angewendet:

Streuung zwischen den Gruppen der Werte auf Diskriminanzachse

$$B_d = \sum_{j=1}^{g} n_j(\bar{d}_j-\bar{d})^2. \tag{2-8}$$

Streuung innerhalb der Gruppen

$$W_d = \sum_{j=1}^{g} \sum_{i=1}^{n_j} (d_{ji} - \bar{d}_j)^2. \tag{2-9}$$

Das Ziel einer optimalen Trennung kann erreicht werden, wenn das **Diskriminanzkriterium** λ maximiert wird, d.h. $\lambda = \frac{B_d}{W_d} \rightarrow Max!$

$$\lambda = \frac{B_d}{W_d} = \frac{\sum_{j=1}^{g} n_j \cdot (\bar{d}_j - \bar{d})^2}{\sum_{j=1}^{g} \sum_{i=1}^{n_j} \cdot (\bar{d}_{ji} - \bar{d}_j)^2} \tag{2-10}$$

2.3.2 Mathematischer Hintergrund

a. Notwendige Erweiterung der Notation
b. Bilden der Streumatrizen (Covarianzmatrixen) **B** und **W** zu den Varianzen von B_d und W_d.
c. Umformung von $B_d = \sum_{j=1}^{g} n_g (\bar{d}_j - \bar{d})^2 = \mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a}$
d. Umformung von $W_d = \sum_{j=1}^{g} \sum_{i=1}^{n_j} (d_{ji} - \bar{d}_j)^2 = \mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}$
e. Maximierung von λ
f. Berechnung der Diskriminanzkoeffizienten

2.3.2.a Notwendige Erweiterung der Notation

Darstellung der Diskriminanzfunktion in Matrizenform

Für die Berechnung der Streumatrizen und zur Darstellung der Diskriminanzfunktion ist es sinnvoll, den Datensatz **X** und die Diskriminanzkoeffizienten **a** in Matrizenform resp. als Vektoren darzustellen:

$$\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}, \mathbf{a_0} = \begin{pmatrix} a_0 \\ a_0 \\ \vdots \\ a_0 \end{pmatrix}, \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, \mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{pmatrix}. \tag{2-11}$$

Für n Untersuchungsobjekte in einer Gruppe und m Merkmalsvariablen resultiert eine (n x m) Matrix **X** mit n Zeilen und m Spalten.

Aus der Diskriminanzfunktion berechnen sich dann allgemein die Diskriminanzwerte:

$$\mathbf{d} = \mathbf{a_0} + \mathbf{X} \cdot \mathbf{a}, \tag{2-12a}$$

resp. für eine Gruppe j:

$$\mathbf{d_j} = \mathbf{a_0} + \mathbf{X_j} \cdot \mathbf{a}. \qquad \text{(2-12b)}$$

Das i-te Objekt der j-ten Gruppe berechnet sich also:

$$d_{ji} = a_0 + a_1 x_{ji1} + a_2 x_{ji2} + ... + a_m x_{jim} \qquad \text{(2-13)}$$

Merkmalsvektor

Der Vektor $\mathbf{x_{ji}}$ enthält als Elemente die $k = 1, ..., m$ Merkmale der i-ten Beobachtung einer Gruppe j:

$$\mathbf{x_{ji}} = \begin{pmatrix} x_{ji1} \\ x_{ji2} \\ \vdots \\ x_{jim} \end{pmatrix} \qquad \text{(2-14)}$$

Darstellung der Mittelwerte in Vektorform

Diese Notation betrifft immer Beobachtungen und eine Gruppe, und es wird dann innerhalb und über Gruppen summiert. Zeilen von $\bar{x}_j$ enthalten dann die Werte des k-ten Merkmales der j-ten Gruppe. Zeilen von $\bar{x}$ enthalten die Mittelwerte des k-ten Merkmales über alle Gruppen.

Für $j = 1$ und $k = 2$ schreibt man beispielsweise:

Gruppenmittelwerte Gesamtmittelwerte

$$\bar{\mathbf{x}}_1 = \begin{pmatrix} \bar{\mathbf{x}}_{11} & G_1, X_1 \\ \bar{\mathbf{x}}_{12} & G_1, X_2 \end{pmatrix} \quad \text{resp.} \quad \bar{\mathbf{x}} = \begin{pmatrix} \bar{\mathbf{x}}_1 & \text{alle } G, X_1 \\ \bar{\mathbf{x}}_2 & \text{alle } G, X_2 \end{pmatrix} \qquad \text{(2-15)}$$

2.3.2.b Bilden der Streumatrizen (Covarianzmatrizen) **B** und **W** zu den Varianzen von B_d und W_d.

Streumatrix **B**

Bei 2 Merkmalen ergibt sich eine (2 x 2) - Streumatrix, allgemein ist sie vom Typ (k x k).

Die Berechnung der Streumatrix **B** der Zwischenquadratsummen für eine Gruppe ergibt sich aus dem dyadischen Produkt der Differenz zwischen dem Vektor der Gruppenmittelwerte und dem Vektor der Gesamtmittelwerte, dann über alle Gruppen summiert. Unter Verwendung der unter (2-15) eingeführten Notation finden wir für die Gruppe 1:

$$n_1 \cdot (\mathbf{x_1} - \bar{\mathbf{x}})(\mathbf{x_1} - \bar{\mathbf{x}})' = n_1 \cdot \begin{pmatrix} \bar{x}_{11} - \bar{x}_1 \\ \bar{x}_{12} - \bar{x}_2 \end{pmatrix} \begin{pmatrix} \bar{x}_{11} - \bar{x}_1 & \bar{x}_{12} - \bar{x}_2 \end{pmatrix}'$$

$$= \begin{bmatrix} n_1 \cdot (\bar{x}_{11} - \bar{x}_1)^2 & n_1 \cdot (\bar{x}_{11} - \bar{x}_1)(\bar{x}_{12} - \bar{x}_2) \\ n_1 \cdot (\bar{x}_{12} - \bar{x}_2)(\bar{x}_{11} - \bar{x}_1) & n_1 \cdot (\bar{x}_{12} - \bar{x}_2)^2 \end{bmatrix} \qquad \text{(2-16)}$$

Bei mehreren Gruppen, d.h. $j \geq 2$, werden die Elemente der Matrix über die Gruppen summiert. Mit k = 2 Merkmalen hat die Streumatrix **B** zwischen den Gruppen folgende Gestalt:

$$\mathbf{B} = \sum_{j=1}^{g} n_j(\mathbf{x_j} - \bar{\mathbf{x}})(\mathbf{x_j} - \bar{\mathbf{x}})', \text{ resp. in Matrizenform:}$$

$$\mathbf{B} = \begin{bmatrix} \sum_{j=1}^{g} n_j \cdot (\bar{x}_{j1} - \bar{x}_1)^2 & \sum_{j=1}^{g} n_g \cdot (\bar{x}_{j1} - \bar{x}_1)(\bar{x}_{j2} - \bar{x}_2) \\ \sum_{j=1}^{g} n_j \cdot (\bar{x}_{j2} - \bar{x}_2)(\bar{x}_{j1} - \bar{x}_1) & \sum_{j=1}^{g} n_j \cdot (\bar{x}_{j2} - \bar{x}_2)^2 \end{bmatrix} \tag{2-17}$$

Streumatrix **W**

Die Streumatrix **W** wird innerhalb der Gruppen gebildet. Die Berechnung für eine Gruppe erfolgt als dyadisches Produkt der Differenz zwischen Merkmalsvektor (2-14) und dem Vektor der Gruppenmittelwerte (2-15), dann über alle Gruppen summiert:

$$\mathbf{W} = \sum_{j=1}^{g} \sum_{i=1}^{n_j} (\mathbf{x_{ji}} - \bar{\mathbf{x}}_\mathbf{j})(\mathbf{x_{ji}} - \bar{\mathbf{x}}_\mathbf{j})', \text{ resp. in Matrizenform:}$$

$$\mathbf{W} = \begin{bmatrix} \sum_{j=1}^{g} \sum_{i=1}^{n_j} (x_{ji1} - \bar{x}_{j1})^2 & \sum_{j=1}^{g} \sum_{i=1}^{n_j} (x_{ji1} - \bar{x}_1)(x_{ji2} - \bar{x}_{j2}) \\ \sum_{j=1}^{g} \sum_{i=1}^{n_j} (x_{ji2} - \bar{x}_{j2})(x_{ji1} - \bar{x}_{j1}) & \sum_{j=1}^{g} \sum_{i=1}^{n_j} (x_{ji2} - \bar{x}_{j2})^2 \end{bmatrix} \tag{2-18}$$

2.3.2.c Umformung von $B_d = \sum_{j=1}^{g} n_j(\bar{d}_j - \bar{d})^2 = \mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a}$

Der Gruppenmittelwert auf der Diskriminanzachse berechnet sich aus der Diskriminanzfunktion wie folgt:

$$\bar{d}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} d_{ji} \;=\; \frac{1}{n_j} \sum_{i=1}^{n_j} \mathbf{a}'\mathbf{x_{ji}} = \mathbf{a}' \frac{1}{n_j} \sum_{i=1}^{n_i} \mathbf{x_{ji}} \;=\; \mathbf{a}' \cdot \bar{\mathbf{x}}_\mathbf{j}. \tag{2-19}$$

Der Vektor $\bar{\mathbf{x}}_\mathbf{j}$ bildet die Koordinaten des Gruppenzentroids im X-Raum
d.h. $\bar{d}_j$ = Abbildung von $\bar{\mathbf{x}}_\mathbf{j}$ auf d.

Der Gesamtmittelwert auf der Diskriminanzachse berechnet sich aus der Diskriminanzfunktion wie folgt:

$$\bar{d} = \frac{1}{\sum_{j=1}^{g} n_j} \sum_{j=1}^{g} \sum_{i=1}^{n_j} d_{ji} \;=\; \frac{1}{\sum_{j=1}^{g} n_j} \sum_{j=1}^{g} \sum_{i=1}^{n_j} \mathbf{a}'\mathbf{x_{ji}} = \mathbf{a}' \frac{1}{\sum_{j=1}^{g} n_j} \sum_{j=1}^{g} \sum_{i=1}^{n_i} \mathbf{x_{ji}} \;=\; \mathbf{a}' \cdot \bar{\mathbf{x}}. \tag{2-20}$$

Der Vektor $\bar{\mathbf{x}}$ bildet die Koordinaten des Gesamtzentroids im X-Raum.
d.h. $\bar{d}$ = Abbildung von $\bar{\mathbf{x}}$ auf d.

Unter Anwendung von (2-19) und (2-20) kann B_d nun notiert werden als:

$$B_d = \sum_{j=1}^{g} n_j(\bar{d}_j - \bar{d})^2 = \sum_{j=1}^{g} n_j(\mathbf{a}'\bar{\mathbf{x}}_\mathbf{j} - \mathbf{a}'\bar{\mathbf{x}})^2 = \sum_{j=1}^{g} n_j\{\underbrace{\mathbf{a}'(\bar{\mathbf{x}}_\mathbf{j} - \bar{\mathbf{x}})}_{b}\}^2 = \sum_{j=1}^{g} n_j \cdot b \cdot b'$$

$$= \sum_{j=1}^{g} n_j \cdot \underbrace{\mathbf{a}'(\bar{\mathbf{x}}_\mathbf{j} - \bar{\mathbf{x}})}_{b} \underbrace{(\bar{\mathbf{x}}_\mathbf{j} - \bar{\mathbf{x}})'\mathbf{a}}_{b'} = \mathbf{a}' \underbrace{\sum_{j=1}^{g} n_j(\bar{\mathbf{x}}_\mathbf{j} - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\mathbf{j} - \bar{\mathbf{x}})'}_{\mathbf{B}} \mathbf{a} = \mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a}$$

$$\rightarrow \; B_d = \sum_{j=1}^{g} n_j(\bar{d}_j - \bar{d})^2 = \mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a} \quad \text{q.e.d.} \tag{2-21}$$

2.3.2.d Umformung von $W_d = \sum_{n=1}^{g} \sum_{i=1}^{n_g} (d_{ji} - \bar{d}_j)^2 = \mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}$

Unter Anwendung von (2-20) kann W_d nun notiert werden als:

$$W_d = \sum_{j=1}^{g} \sum_{i=1}^{n_j} (d_{ji} - \bar{d}_j)^2 = \sum_{j=1}^{g} \sum_{i=1}^{n_j} (\mathbf{a}'\mathbf{x_{ji}} - \mathbf{a}'\bar{\mathbf{x}}_\mathbf{j})^2 = \sum_{j=1}^{g} \sum_{i=1}^{n_j} \{\underbrace{\mathbf{a}'(\mathbf{x_{ji}} - \bar{\mathbf{x}}_\mathbf{j})}_{b}\}^2$$

$$= \sum_{j=1}^{g} \sum_{i=1}^{n_j} b \cdot b' = \sum_{j=1}^{g} \sum_{i=1}^{n_j} \underbrace{\mathbf{a}'(\mathbf{x_{ji}} - \bar{\mathbf{x}}_\mathbf{j})}_{b} \underbrace{(\mathbf{x_{gi}} - \bar{\mathbf{x}}_\mathbf{j})'\mathbf{a}}_{b'} = \mathbf{a}' \underbrace{\sum_{j=1}^{g} \sum_{i=1}^{n_j} (\mathbf{x_{ji}} - \bar{\mathbf{x}}_\mathbf{j})(\mathbf{x_{ji}} - \bar{\mathbf{x}}_\mathbf{j})'}_{\mathbf{W}} \mathbf{a}$$

$$\rightarrow \; W_d = \sum_{n=1}^{g} \sum_{i=1}^{n_g} (d_{ji} - \bar{d}_j)^2 \;=\; \mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a} \quad \text{q.e.d.} \tag{2-22}$$

2.3.2.e Maximierung von λ

Die unter 2.3.2.c und 2.3.2.d hergeleiteten Umformungen erlauben nun die Maximierung des Diskriminanzkriteriums λ:

$$\lambda = \frac{\mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a}}{\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}} \quad \rightarrow max! \tag{2-23}$$

Das **Diskriminanzkriterium** λ ist im Hinblick auf die **Diskriminanzkoeffizienten** $\mathbf{a_0}$, $\mathbf{a_1}, \ldots$, $\mathbf{a_m}$ so zu maximieren, d.h. die Diskriminanzkoeffizienten sind so zu bestimmen, dass das Verhältnis aus der Streuung zwischen den Gruppen (B_d) und der Streuung innerhalb der Gruppen (W_d) maximal wird.

λ wird maximal, wenn die 1. Ableitung (Gradient) von $\lambda(\mathbf{a})$, die einen Vektor ergibt, gleich dem Nullvektor gesetzt wird:

$$\frac{\partial}{\partial \mathbf{a}}(\lambda) = \frac{\partial}{\partial \mathbf{a}}\left[\frac{\mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a}}{\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}}\right] = \mathbf{0} \tag{2-24}$$

Gemäss Quotientenregel folgt:

$$\left[\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}\right] \cdot \frac{\partial}{\partial \mathbf{a}}\left[\mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a}\right] - \left[\mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a}\right] \cdot \frac{\partial}{\partial \mathbf{a}}\left[\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}\right] = \mathbf{0} \tag{2-25}$$

Die verbleibenden Ableitungen erfolgen nach den Regeln der Ableitung des Rayleigh-Quotienten nach dem Satz von Courant-Fischer (siehe Appendix B). Appendix B
Man erhält:

$$\frac{\partial}{\partial \mathbf{a}}\left[\mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a}\right] = 2\mathbf{Ba}, \text{ resp. } \frac{\partial}{\partial \mathbf{a}}\left[\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}\right] = 2\mathbf{Wa} \tag{2-26}$$

Der Ausdruck für die Maximierung von (2-25) vereinfacht sich somit zu:

$$\left[\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}]2\mathbf{Ba}\right] - \left[\mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a}\right] 2\mathbf{Wa} = \mathbf{0} \tag{2-27}$$

Division durch $[\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}]$ ergibt

$$\frac{[\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}]}{[\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}]}\mathbf{Ba} - \underbrace{\frac{[\mathbf{a}' \cdot \mathbf{B} \cdot \mathbf{a}]}{[\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}]}}_{\lambda}\mathbf{Wa} = \mathbf{0} \tag{2-28}$$

Durch Verwendung der Definition von λ (2-23) erhalten wir weiter durch Umformen ein Eigenwertproblem:

$$\mathbf{Ba} - \lambda\mathbf{Wa} = \mathbf{0}0 \quad \rightarrow \quad | \quad \cdot\mathbf{W}^{-\mathbf{1}}$$

$$\rightarrow \quad \underbrace{\mathbf{W}^{-\mathbf{1}}\mathbf{B}}_{\text{Eigenmatrix}} \ \mathbf{a} - \lambda \underbrace{\mathbf{W}^{-\mathbf{1}}\mathbf{W}}_{\mathbf{I}} \cdot \mathbf{a} = \mathbf{0}$$

$$\left(\mathbf{W}^{-\mathbf{1}}\mathbf{B} - \lambda \cdot \mathbf{I}\right) \cdot \mathbf{a} = \mathbf{0} \tag{2-29}$$

Die Lösung dieses allg. Eigenwertproblemes ergibt **a** als Eigenvektor zum grössten Eigenwert von $[\mathbf{W}^{-\mathbf{1}}\mathbf{B}]$.

$\Longrightarrow$ λ = Eigenwert des Eigenwertproblems

a = dazugehöriger Eigenvektor.

2.3.2.f Berechnung der Diskriminanzkoeffizienten

Aus Gleichung (2-29) können für nicht-triviale Lösungen, d.h. $\mathbf{a} \neq \mathbf{0}$, die Eigenwerte λ bestimmt werden. Bedingung ist, dass die Determinante dieser Gleichung 0 sein muss, also:

$$|\,\mathbf{W}^{-\mathbf{1}}\mathbf{B} - \lambda\cdot\mathbf{I}\,| = 0 \tag{2-30}$$

Einsetzen des grösstem Eigenwerts in Gleichung (2-30) bestimmt den zugehörigen Eigenvektor.

In der Regel wird dieser Eigenvektor auf 1 normiert, so dass gilt:

$$\mathbf{a}'\mathbf{a} = \sum_{=1}^{m} a_j^2 = \mathbf{a_1^2} + \mathbf{a_2^2} + \ldots + \mathbf{a_m^2} = \mathbf{1} \tag{2-31}$$

Diese Art der Normierung ist im Falle der LDA ungünstig und man wählt eine andere, wie im anschliessenden Kapitel 2.4 gezeigt wird.

2.4 Normierung der Koeffizienten

Anstelle der üblichen Normierung wird für Diskriminanzanalyse eine andere Art der Normierung gewählt. Hier soll die gepoolte Varianz der Diskriminanzvariablen 1 sein:

$$\left(s_d^2\right)^{\text{pool}} \stackrel{!}{=} 1 \tag{2-32}$$

Die Streuung innerhalb der Gruppen $W_d = \sum_{j=1}^{g}\sum_{i=1}^{n_g}(d_{ji} - \bar{d}_j)^2$ der Diskriminanzwerte (2-9) als $\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}$ wurde in Kapitel 2.3.2.d hergeleitet:

Gefordert wird nun mit n = $\sum_{j=1}^{g} n_j$:

$$\left(s_d^2\right)^{\text{pool}} = \frac{1}{n-g} \cdot W_d \stackrel{!}{=} 1 \tag{2-33}$$

Daraus folgt:

$$W_d = (\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}) = (n-g), \tag{2-34}$$

wobei n die Gesamtzahl der Objekte und g die Anzahl der Gruppen festlegt. Diese

Normierung standardisiert die gepoolte Varianz auf 1. Allfällige Unterschiede der Streuungen zwischen den Gruppen können damit ausgeglichen werden.

Mittels der Transformation von $\mathbf{a}$ auf $\tilde{\mathbf{a}}$ kann die Bedingung von (2-33) erzielt werden. Gesucht ist also die Proportionalitätskonstante γ, welche wir durch einsetzen finden:

$$\gamma^2 \cdot \underbrace{\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}}_{\text{Skalar}} \stackrel{!}{=} n-g \quad | \text{ auflösen nach } \gamma \;\Rightarrow\; \gamma = \sqrt{\frac{n-g}{\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}}}. \tag{2-35}$$

Normierten Diskriminanzkoeffizienten transformiert:

$$\text{normierte Diskriminanzkoeffizienten } \tilde{\mathbf{a}} = \sqrt{\frac{n-g}{\mathbf{a}' \cdot \mathbf{W} \cdot \mathbf{a}}} \cdot \mathbf{a} \tag{2-36}$$

Für die Ermittlung der normierten Diskriminanzkoeffizienten wird lediglich ein bestehender Lösungvektor $\mathbf{a}$ mit einer Konstanten γ multipliziert. D.h. diese Transformation wirkt sich <u>nicht</u> auf Lage, sondern lediglich auf Skalierung der d-Werte aus.

2.5 Festlegen der definitiven Diskriminanzfunktion

Konstante $\mathbf{a_0}$ wird so bestimmt, dass der Gesamtmittelwert der Diskriminanzwerte 0 wird $\rightarrow$ $\bar{d} \stackrel{!}{=} 0$

Notation (2-12a) der Diskriminanzfunktion wird damit zu:

$$\mathbf{a_0} + \bar{\mathbf{X}} \cdot \tilde{\mathbf{a}} \stackrel{!}{=} 0 \quad \left((s_d^2)^{\text{pool}} \stackrel{!}{=} 1 \right),$$
$$\Longrightarrow \mathbf{a_0} = -\bar{\mathbf{X}} \cdot \tilde{\mathbf{a}} \tag{2-37}$$

Damit wird der kritische Diskriminanzwert $\mathbf{d_c} = 0$. Im weiteren erhält man dadurch die Konstante $\mathbf{a_0}$ sowie die normierten Diskriminanzwerte $\tilde{\mathbf{a}}_\mathbf{j}$ $(j = 1, \ldots, m)$ mit dem Mittelwert 0 und der gepoolten Varianz von eins.Die Diskriminazwerte werden jetzt in Einheiten der Standardabweichung von den Gruppenzentroiden gemessen.

Die **Die normierte Diskriminanzfunktion** lautet somit:

$$\mathbf{d} = \mathbf{a_0} + \mathbf{X} \cdot \tilde{\mathbf{a}} \tag{2-38}$$

Konstante $\mathbf{a_0}$ hat keinen Einfluss auf Streuung der Diskriminazwerte. Sie bewirkt lediglich eine Skalenverschiebung der Diskriminanzwerte.

2.6 Standardisierte Diskriminanzfunktion

Normierte Diskriminanzkoeffizienten werden durch Masseinheiten und Streuungen der Variablen beeinflusst. Zur Ermittlung der Bedeutungsrangfolge ist die Ausschaltung dieser Dimensionen vorteilhaft, was durch eine **Standardisierung der Diskriminanzkoeffizienten** erfolgt.

standardisierter Diskrimininanzkoeffizient = a_j^*

$a_j^* \ = \ \tilde{a} \cdot s_j^{\text{pool}}$ | SD von X_j über beide Gruppen.

$$s_j^{\text{pool}} \ = \ \frac{1}{\sqrt{n-2}} w_{jj}^{1/2}$$

in Matrizenschreibweise: $\mathbf{a}^* \ = \dfrac{1}{\sqrt{n-2}} \text{diag}(\mathbf{W}^{1/2}) \cdot \tilde{a}$

Relative Bedeutung der einzelnen Variablen X_j:

$$PT_j \ = \ \left(\left| a_j^* \right| \Big/ \sum_{k=1}^{m} |a_k^*| \right) \cdot 100\%$$

3 Referenzen

[1] www.uni-kassel.de//fb07/ivwl/Lehrveranstaltungen/Multivariate14_Diskriminanzanalyse1_%20(5).pdf, download 11.11.2022
[2] www.sjsu.edu/faculty/guangliang.chen/Math253S20/lec4RayleighQuotient.pdf, download 11.11.2022
[3] www.uwe-mortensen.de/KlassifDiskriminanzAllgemein.pdf, download 11.11.2022

Appendices

A Varianzanalyse Anova

Notation der Variablen und Indizes

Grösse	*Abk.*	*Index*	*max.Wert*	*Bereich*
Gruppe	G	j	g	$1 \leq j \leq g$
Objekt	O	i	n_j	$1 \leq i \leq n_j$
Merkmal	X	k	m	$1 \leq k \leq m$

A.1 Einleitung

Das Prinzip der Varianzanalyse (ANalysis Of VAriance)) besteht in der Zerlegung der Varianz einer metrisch skalierten abhängigen Variablen. Die Gesamtvarianz T_S setzt sich aus der sogenannten ≪Varianz innerhalb (within, nicht erklärter Anteil) der Gruppen≫, W_S, und der ≪Varianz zwischen (between, erklärter Anteil) den Gruppen≫, B_S, zusammen. Das Verhältnis F der beiden Anteile wird im Rahmen einer Varianzanalyse in einem F-Test verglichen. Wenn keine Gruppenunterschiede bestehen, kann die Varianz zwischen den Gruppen nicht von der Varianz innerhalb der Gruppen unterschieden werden (kleines F, Nullhypothese nicht verworfen). Ist die Varianz zwischen den Gruppen markant grösser als die Varianz innerhalb der Gruppen, lassen sich die Effekte der abhängigen Variablen signifikant nachweisen (grosses F, Nullhypothese verworfen).

A.2 Notation der Anova

Gegeben sei eine Stichprobe mit gesamthaft N beobachteten Objekten O (Beobachtung, Messwerte; normalverteilt) mit unabhängigen Eigenschaften A, Faktoren genannt (einfaktoriell, zweifaktoriell, etc.). Jeder Faktor sei in Stufen unterteilt, jede Stufe (Faktorstufe) entspricht somit einer Gruppe des Faktors.

Wir beschränken uns im Weiteren auf den einfaktoriellen Fall, d.h. wir betrachten nur eine unabhänge Eigenschaft A mit verschiedenen Stufen (Gruppen).

N setzt sich zusammen aus der Summe aller beobachteten Objekten über alle Gruppen, wobei n_j jeweils den Umfang der j-ten Gruppe eines Faktors darstellt.

$$N = \sum_{j=1}^{g} n_j \ . \tag{A-1}$$

Der Wert einer Beobachtung sei:
x_{ji}, wobei:

Index i: i-tes Objekt (Untersuchungseinheit): i = 1, 2,...,n_j

Index j: j-te Gruppe (Faktorstufe A_j): j = 1, 2,...,g

Dann berechnet sich der Mittelwert innerhalb der Gruppe j durch:

$$\bar{x}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ji} , \tag{A-2}$$

und der Gesamtmittelwert durch:

$$\bar{x} = \frac{1}{N} \sum_{j=1}^{g} \sum_{i=1}^{n_j} x_{ji} = \frac{1}{N} \sum_{j=1}^{g} n_j \bar{x}_j . \tag{A-3}$$

Für <u>einen</u> Faktor mit mehreren Stufen (Gruppen) gilt allgemein die Varianzzerlegung,

$$\sum_{j=1}^{g} \sum_{i=1}^{n_j} (x_{ji} - \bar{x})^2 = \sum_{j=1}^{g} n_j (\bar{x}_j - \bar{x})^2 + \sum_{j=1}^{g} \sum_{i=1}^{n_j} (x_{ji} - \bar{x}_j)^2 , \tag{A-4}$$

wie weiter unten noch bewiesen wird, abgekürzt oft geschrieben als:

$$T_S \;=\; B_S \;+\; W_S , \tag{A-5}$$

mit folgenden Bedeutungen:

T_S : <u>Totale</u> Streuung. Abweichungen der Objekte vom Gesamtmittelwert $\bar{x}$:

$$T_S = \sum_{j=1}^{g} \sum_{i=1}^{n_j} (x_{ji} - \bar{x})^2 , \tag{A-6}$$

B_S: Streuung <u>zwischen den Gruppen (between)</u>, erklärte Abweichungen:

$$B_S = \sum_{j=1}^{g} n_j (\bar{x}_j - \bar{x})^2 , \tag{A-7}$$

W_S: Streuung <u>innerhalb der Gruppen (within)</u>, nicht erklärte Abweichungen:

$$W_S = \sum_{j=1}^{g} \sum_{i=1}^{n_j} (x_{ji} - \bar{x}_j)^2 . \tag{A-8}$$

Das Verhältnis F der beiden Anteile B_S / W_S wird im Rahmen einer Varianzanalyse in einem F-Test verglichen, um zu prüfen, ob sich Effekte der abhängigen Variablen signifikant nachweisen lassen. Dieses Verhältnis gibt deshalb wichtige Hinweise darüber wie die Hypothesen zu bewerten sind,

$$F = \frac{B_S}{W_S} = \frac{\sum_{j=1}^{g} n_j \cdot (\bar{x}_j - \bar{x})^2}{\sum_{j=1}^{g} \sum_{i=1}^{n_j} \cdot (\bar{x}_{ji} - \bar{x}_j)^2} . \tag{A-9}$$

A.3 Beweis von $T_S = B_S + W_S$

Lemma: Es seien $a_1, ...a_n$ beliebige n reelle Zahlen und $\bar{a} = \sum_{i=1}^{n} a_i/n$ sei ihr arithmetisches Mittel. Weiter sei b eine andere beliebige reelle Zahl. Dann gilt:

$$\sum_{i=1}^{n}(a_i - \bar{a})^2 = \sum_{i=1}^{n}(a_i - b)^2 - n(\bar{a} - b)^2 \ . \tag{A-10}$$

Beweis
Es gilt:

$$\begin{aligned}\sum_{i=1}^{n}(a_i - \bar{a})^2 &= \sum_{i=1}^{n}(a_i - b + b - \bar{a})^2 = \sum_{i=1}^{n}((a_i - b) - (\bar{a} - b))^2 \\ &= \sum_{i=1}^{n}(a_i - b)^2 + n(\bar{a} - b)^2 - 2(\bar{a} - b)\sum_{i=1}^{n}(a_i - b)\end{aligned} \tag{A-11}$$

Da

$$\begin{aligned}\sum_{i=1}^{n}(a_i - b) &= a_1 - b + a_2 - b, ... + a_n - b \\ &= a_1 + a_2 + ... + a_n \ - \ n \cdot b\end{aligned} \tag{A-12}$$

und

$$\sum_{i=1}^{n} a_i = n \cdot \bar{a} \ ,$$

folgt

$$\sum_{i=1}^{n}(a_i - \bar{a})^2 = \sum_{i=1}^{n}(a_i - b)^2 - n(\bar{a} - b)^2 \ . \qquad q.e.d. \tag{A-13}$$

Zerlegung der Quadratsumme

Die Quadratsumme ist über die Gesamtvarianz definiert. Weiter sei n_j der Stichprobenumfang der j-ten Gruppe und $N = n_1 + n_2 + ... + n_j$. Der Mittelwert innerhalb einer Gruppe $\bar{x}_j$ wurde in (A-2) definiert, entsprechend beschreibt (A-3) den Gesamtmittelwert $\bar{x}$. Dann kann die Quadratsumme über die Gesamtvarianz

$$QS_{ges} = \sum_{j=1}^{g}\sum_{i=1}^{n_j}(x_{ij} - \bar{x})^2 \tag{A-14}$$

zerlegt werden, und zwar gemäss (A-4),

$$\sum_{j=1}^{g}\sum_{i=1}^{n_j}(x_{ji} - \bar{x})^2 = \sum_{j=1}^{g} n_j(\bar{x}_j - \bar{x})^2 + \sum_{j=1}^{g}\sum_{i=1}^{n_j}(x_{ij} - \bar{x}_j)^2 \ .$$

Beweis: Wir betrachten die j-te Teilsumme

$$\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)^2 \ . \tag{A-15}$$

Anhand des obigen Lemmas erhält man mit $b=\bar{x}_j$ sofort

$$\sum_{i=1}^{n_j}(x_{ji}-\bar{x})^2=\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)^2-n_j(\bar{x}-\bar{x}_j)^2 \ . \tag{A-16}$$

Durch die Umkehr des Vorzeichens im letzten Glied von (A-16) und Summation beider Seiten über die Gruppen$j, \ j=1,...,g$ erhält man die in (A-4) postulierte Zerlegung der Quadratsumme QS_{ges} von (A-14). Die erste Summe auf der rechten Seite entspricht der Quadratsumme zwischen den Gruppen und ist identisch mit (A-7).

$$QS_{zw}=\sum_{j=1}^{g}n_j(x_j-\bar{x})^2 \ . \tag{A-17}$$

Die zweite Summe auf der rechten Seite heisst oft QS_{inn}, entspricht der Quadratsumme innerhalb der Gruppe und ist identisch mit (A-8)

$$QS_{inn}=\sum_{j=1}^{g}\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)^2 \ . \tag{A-18}$$

Symbolisch schreibt man

$$QS_{ges} \ = \ QS_{zw} \ + \ QS_{inn} \equiv T_S \ = \ B_S \ + \ W_S \ . \tag{A-19}$$

B Beweis des Satzes von Courant-Fischer

Nach Vorlesungs-Skripts: G. Chen, San José State University, USA [2]
U. Mortensen, 10.12.2019, Universität Münster, BRD [3]

Der Raleigh-Quotient ist wie folgt definiert :

$$Q(\mathbf{x}) = \frac{\mathbf{x^T A x}}{\mathbf{x^T x}} \tag{B-1}$$

Dieser Ausdruck spielt in der LDA eine zentrale Rolle. Er gilt für alle symmetrischen Matrizen $\mathbf{A} \in \mathbb{R}^{n \times n}$ und ist damit anwendbar zur Lösung des Problems:

$$\max_{(\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = \mathbf{1})} \frac{\mathbf{x^T A x}}{\mathbf{x^T x}} = \lambda_{\max} \tag{B-2}$$

mit $\lambda_{\max}$ als grösster Eigenwert, produziert von $\mathbf{x}$, welches den grössten Eigenwert von $\mathbf{A}$ erzeugt.

Da $Q(\mathbf{x})$ auf Skalierung invariant ist, kann auf die Einheitskugel fokussiert werden:

$$\max_{(\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = \mathbf{1})} \mathbf{x^T A x} = \lambda_{\max} \tag{B-3}$$

Der Satz von Courant-Fischer stellt die Eigenwerte einer symmetrischen oder hermiteschen Matrix als minimale resp. maximale Rayleigh-Quotienten dar.

Den Beweis des Maximierungsproblems erbringen wir mittels zweier Verfahren, nämlich:

- Beweis mit linearer Algebra (Kapitel B.1)

- Beweis mit Differentialrechnung (Kapitel B.2)

B.1 Beweis mit linearer Algebra

Dieses Vorgehen führt über die Spektralzerlegung einer Matrix.

Beweis: $\mathbf{A}$ ist nach Voraussetzung symmetrisch, daher gilt:

$$\mathbf{A} = \mathbf{Q} \Lambda \mathbf{Q^T} \tag{B-4}$$

wobei $\mathbf{Q} = [\mathbf{q_1}, ..., \mathbf{q_n}]$ die Matrix mit den orthonormalen Eigenvektoren,

und $\Lambda = \text{diag}(\lambda_1, ... \lambda_n)$, die Diagonalmatrix mit den entsprechenden Eigenwerten, sortiert in absteigender Reihenfolge (d.h. $\lambda_1 > ... > \lambda_n$), darstellt.

$\mathbf{A}$ besitzt die Eigenwerte λ_i und die korrespondierenden Eigenvektoren $\mathbf{v}_i$. $\mathbf{A}$ und Λ haben die gleichen Eigenwerte. Die Eigenvektoren von Λ sind die Einheitsvektoren $\mathbf{e}_i$ der Standardbasis. Weil $\mathbf{Q}$ orthonormal ist, gilt $\|\mathbf{Qx}\| = \|\mathbf{x}\|$. $\mathbf{x}$ ist orthogonal zu $\mathbf{v}_i$, wenn $\mathbf{Qx}$ orthogonal zum i-ten Einheitsvektor $\mathbf{e}$ ist. Deshalb genügt es, nur den Fall zu betrachten, bei dem $\mathbf{A} = \Lambda$ ist.(alternativ: Mit nachstehender Umformung kann dieser günstige Fall zur Berechnung herbeigeführt werden).

Dann gilt für jeden Einheitsvektor $\mathbf{x}$:

$$\mathbf{x^T A x} \;=\; \mathbf{x^T}(\mathbf{Q}\,\Lambda\,\mathbf{Q^T})\mathbf{x} \;=\; (\mathbf{x^T Q})\,\Lambda\,(\mathbf{Q^T x}) \;=\; \mathbf{y^T}\Lambda\,\mathbf{y} \tag{B-5}$$

Dabei haben wir $\mathbf{y} = \mathbf{Q^T}\,\mathbf{x}$ gesetzt, wobei $\mathbf{y}$, wie (B-6) zeigt, ebenfalls einen Einheitsvektor darstellt.

$$\|\mathbf{y}\|^{\mathbf{2}} \;=\; \mathbf{y^T y} \;=\; (\mathbf{Q^T x})^{\mathbf{T}}(\mathbf{Q^T x}) \;=\; \mathbf{x^T}\underbrace{\mathbf{QQ}^T}_{\mathbf{I}}\mathbf{x} \;=\; \mathbf{x^T x} \;=\; 1 \tag{B-6}$$

Das ursprüngliche Optimierungsproblem (B-3) lautet nun unter Berücksichtigung von (B-5) folgendermassen:

$$\max_{(\mathbf{x}\in\mathbb{R}^n:\|\mathbf{x}\|=\mathbf{1})} \mathbf{y^T}\underbrace{\Lambda}_{\textbf{diagonal}}\mathbf{y}. \tag{B-7}$$

Um dieses neu formulierte (vereinfachte) Problem zu lösen, schreiben wir für den Vektor $\mathbf{y} = (y_1 ...., y_n)^T$:

$$\mathbf{y^T \Lambda y} = [y_1,...,y_n]\begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \sum_{i=1}^{n}\lambda_i y_i^2. \tag{B-8}$$

Die Eigenwerte sind in absteigender Reihenfolge sortiert und es gilt:

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n. \tag{B-9}$$

Wegen (B-5) muss gelten, dass:

$$y_1^2 + y_2^2 + ... + y_n^2 = 1 \tag{B-10}$$

Unter Berücksichtigung von (B-9) und (B-10) erreicht (B-8) seinen Maximalwert dann, wenn wir den grössten Eigenwert mit dem entsprechenden Einheitsvektor multiplizieren;

$$\begin{aligned} &\lambda_{\max} = \lambda_1, \text{sowie} \\ &y_1^2 = 1,\; y_2^2 = ... = y_n^2 = 0,\; \text{d.h.} \\ &\mathbf{y} = \mathbf{e_1}\text{verwenden.} \end{aligned} \tag{B-11}$$

Für den grössten Wert von (B-8) gilt also:

$$\max_{(\mathbf{y^T \Lambda y})} = \lambda_1. \tag{B-12}$$

Den Maximierer des zugehörigen Originalvektors $\mathbf{x}$ erhalten wir unter verwendung der Substitution aus (B-5):

$$\mathbf{x} = \mathbf{Qy} = \mathbf{Q}(\pm\mathbf{e_1}) = \pm\mathbf{q_1} \tag{B-13}$$

Daraus folgt, dass die quadratische Form $\mathbf{x^T A x}$ ihr Maximum mit dem grössten Eigenvektor $\pm\mathbf{q_1}$ und dem grössten Eigenwert von $\mathbf{A}$, $\lambda_1 = \lambda_{\max}$, erreicht.

B.2 Vorgehen mittels Differentialrechnung

Alternativ können wir Theorem (B-3) mittels der Methode des Lagrange Multiplikators beweisen. Die dazugehörende Lagrange-Funktion lautet:

$$\mathbf{L}(\mathbf{x},\lambda) = \mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x} - \underbrace{\lambda(\|\mathbf{x}\|^2 - 1)}_{\text{Randbedingung=0}}. \tag{B-14}$$

Zum Auffinden der Extrema von (B-14) berechnen wir den Gradienten der Funktion und setzen ihn 0.

Zuvor müssen wir aber ein paar nützliche Regeln zur Ableitung von Funktionen mit vektorieller Variablen $\mathbf{x}$ bereitstellen. Es sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ eine konstante symmetrische Matrix, $\mathbf{B} \in \mathbb{R}^{n \times n}$ eine konstante rechteckige Matrix, und $\mathbf{a} \in \mathbb{R}^{n}$ ein konstanter Vektor.

Regeln für ausgewählte Ableitungen:

1. Partielle Ableitung des Produktes eines konstanten Vektors mit einer vektoriellen Variablen:
$$\frac{\partial}{\partial x}(\mathbf{a}^{\mathbf{T}}\mathbf{x}) = \mathbf{a} \tag{B-15}$$
Beweis: Wir nehmen für alle Terme von $1 \leq k \leq n$ die k-te Ableitung, daraus folgt (B-15):
$$\frac{\partial}{\partial x_k}(\mathbf{a}^{\mathbf{T}}\mathbf{x}) = \frac{\partial}{\partial x_k}\left(\sum_{i=1}^{n} a_i x_i\right) = a_k$$

2. Ableitung des Quadrats der Norm einer vektoriellen Variable:
$$\frac{\partial}{\partial x}\left(\|\mathbf{x}\|^2)\right) = 2\mathbf{x} \tag{B-16}$$
Beweis: Wir nehmen für alle Terme von $1 \leq k \leq n$ die k-te partielle Ableitung, daraus folgt (B-16)
$$\frac{\partial}{\partial x}\left(\|\mathbf{x}\|^2\right) = \frac{\partial}{\partial x_k}\left(\sum_{i=1}^{n} x_i^2\right) = 2x_k$$

3. Ableitung einer quadratischen Form:
$$\frac{\partial}{\partial x}\left(\mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x}\right) = 2\mathbf{A}\mathbf{x}. \tag{B-17}$$
Gegeben sei eine konstante symmetrische Matrix $\mathbf{A} \in \mathbb{R}^{(n\ x\ n)}$ sowie Vektoren $\mathbf{x} \in \mathbb{R}^{n}$. Wegen der Symmetrie gilt $\mathbf{A} = \mathbf{A}^{T}$. Dann beschreibt der Ausdruck (B-17) den Gradienten der quadratischen Form $f(\mathbf{x}) = \mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x}$.

Für das Verständnis ist der Teil (B-17.1) bis (B-17.12) nicht notwendig und der Leser kann direkt zum Kapitel 4 ”Ableitung des Normquadrates einer Matrizenmultiplikation” springen (Formel B-18). Der Teil zum Beweis von (B-17) dient dem mathematisch interessierten Leser.

Beweis:

Wir bilden die partiellen Ableitungen in Richtung der generischen Vektorkomponenten $\mathbf{x}_k$, mit $k = 1, ..., n$. Diese n partiellen Ableitungen ergeben zusammengefasst den Lösungsvektor.

Der Beweis ist allgemein formuliert und jeweils am Beispiel der Funktion $f(\mathbf{x}) = \mathbf{x}^\mathbf{T}\mathbf{A}\mathbf{x}$ mit Vektor $\mathbf{x} \in \mathbb{R}^3$ und Matrix $\mathbf{A} \in \mathbb{R}^{(3\ x\ 3)}$, sowie der generischen Vektorkomponente $x_k,\ k = 1$, illustriert.

Allgemein: Die partielle Ableitung der quadratischen Form (B-17), ausgedrückt als Summe, lautet :

$$\frac{\partial}{\partial x_k}\left(\mathbf{x}^\mathbf{T}\mathbf{A}\mathbf{x}\right) = \frac{\partial}{\partial x_k}\left(\sum_i \sum_j a_{ij} x_i x_j\right). \tag{B-17.1}$$

Beachte, dass der Klammerausdruck der rechten Seite das Skalarprodukt darstellt, wobei die Komponenten des Vektors $\mathbf{x}^T$ mit dem Laufindex i und diejenigen des Vektors $\mathbf{x}$ mit dem Laufindex j bezeichnet sind.

Beispiel: Aussummiert lautet das Skalarprodukt $f(\mathbf{x}) = \mathbf{x}^\mathbf{T}\mathbf{A}\mathbf{x}$ folgendermassen:

$$\left(\sum_i \sum_j a_{ij} x_i x_j\right) = \begin{array}{rrr} a_{11}x_1^2 & +\ a_{12}x_1x_2 & +\ a_{13}x_1x_3 \\ +\ a_{21}x_2x_1 & +\ a_{22}x_2x_2 & +\ a_{23}x_2x_3 \\ +\ a_{31}x_3x_1 & +\ a_{32}x_3x_2 & +\ a_{33}x_3^2. \end{array} \tag{B-17.2}$$

Wir wenden uns nun der partiellen Ableitung zu. Ausdruck (B-17.1) kann auch geschrieben werden als:

$$\frac{\partial}{\partial x_k}\left(\sum_i \sum_j a_{ij} x_i x_j\right) = \frac{\partial}{\partial x_k}\left(\sum_{\substack{j\neq k \\ i=k}} a_{kj} x_k x_j + \sum_{\substack{i\neq k \\ j=k}} a_{ik} x_i x_j + a_{kk} x_k^2\right). \tag{B-17.3}$$

Für die partielle Ableitung liefern nur die Ausdrücke mit x_k einen Beitrag, für alle anderen Elemente ist Ableitung 0.

Die 1. Summe entspricht der Ableitung der k-ten Zeile, ohne quadratisches Glied, und ergibt:

$$\frac{\partial}{\partial x_k}\left(\sum_{\substack{j\neq k \\ i=k}} a_{kj} x_k x_j\right) = \sum_{j\neq k} a_{kj} x_j. \tag{B-17.4}$$

Die 2. Summe entspricht der Ableitung der k-ten Spalte, ohne quadratisches Glied, und ergibt:

$$\frac{\partial}{\partial x_k}\left(\sum_{\substack{i\neq k\\ j=k}} a_{ik}x_ix_k\right)=\sum_{i\neq k} a_{ik}x_i. \qquad \text{(B-17.5)}$$

Die 3. Summe entspricht der Ableitung des quadratischen Gliedes, und ergibt:

$$\frac{\partial}{\partial x_k}\left(a_{kk}x_k^2\right) \;=\; 2a_{kk}x_k. \qquad \text{(B-17.6)}$$

Zusammengefasst ergibt die 1. partielle Ableitung:

$$\begin{aligned}\frac{\partial}{\partial x_k}\left(\mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x}\right) \;&=\; \sum_{j\neq k} a_{kj}x_j \;+\; \sum_{i\neq k} a_{ik}x_i \;+\; 2a_{kk}x_k\\ &=\; \sum_{j} a_{kj}x_j \;+\; \sum_{i} x_ia_{ik} \qquad \text{(B-17.7)}\end{aligned}$$

<u>Beispiel:</u> Die Vektorkomponente x_k, $k=1$, die sich aus der partiellen Ableitung von (B-17.2) berechnet, benötigt nur die folgenden Term, alle anderen ergeben 0:

$$a_{11}x_1^2+a_{12}x_1x_2+a_{13}x_1x_3+a_{21}x_2x_1+a_{31}x_3x_1 \qquad \text{(B-17.7a)}$$

Formel (B-17.7) angewandt auf (B-17.7a), und umgeformt in Matrizenschreibweise liefert:

$$\sum_{j=1}^{3} a_{kj}x_j+\sum_{i=1}^{3} x_ia_{ik}=\begin{pmatrix} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3\\ +\,a_{11}x_1 & + & a_{21}x_2 & + & a_{31}x_3. \end{pmatrix} \qquad \text{(B-17.8)}$$

Wie man sich leicht überzeugt, entsprechen im Ausdruck (B-17.7) die Terme a_{kj} der k-ten Zeile, die Terme a_{ik} der k-ten Spalte. Für den Rest des Beweises bedienen wir uns der folgenden Notation. Es soll gelten:

$\mathbf{A}(k,:)=k$-te Zeile von $\mathbf{A}$,

$\mathbf{A}(:,k)=k$-te Spalte von $\mathbf{A}$.

Ausdruck (B-17.7) kann somit umformuliert werden zu:

$$\frac{\partial}{\partial x_k}\left(\mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x}\right)=\mathbf{A}(k,:)\mathbf{x}+\mathbf{x}^T\mathbf{A}(:,k). \qquad \text{(B-17.9)}$$

Nach Voraussetzung ist Matrix $\mathbf{A}$ symmetrisch. Daraus folgt für die partielle Ableitung in Richtung von x_k unmittelbar für (B-17.9):

$$=2\mathbf{A}(k,:)\mathbf{x} \qquad \text{(B-17.10)}$$

Es gilt nun diese partiellen Ableitungen in alle Richtungen der möglichen x_k durchzuführen. Zusammengefasst ergibt sich somit die partielle Ableitung zu:

$$\frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x^T A x}\right) = \begin{bmatrix} \frac{\partial}{\partial \mathbf{x_1}}(\mathbf{x^T A x}) \\ \vdots \\ \frac{\partial}{\partial \mathbf{x_k}}(\mathbf{x^T A x}) \\ \vdots \\ \frac{\partial}{\partial \mathbf{x_n}}(\mathbf{x^T A x}) \end{bmatrix} = \begin{bmatrix} 2\mathbf{A}(1,:)\mathbf{x} \\ \vdots \\ 2\mathbf{A}(k,:)\mathbf{x} \\ \vdots \\ 2\mathbf{A}(n,:)\mathbf{x} \end{bmatrix} = 2\mathbf{Ax}. \quad \text{(B-17.11)}$$

<u>Beispiel:</u>

$$\begin{aligned} \frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x^T A x}\right) &= \frac{\partial}{\partial \mathbf{x}}\left(\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) \\ &= 2 \cdot \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \end{aligned} \quad \text{(B-17.12)}$$

<u>Numerisches Beispiel</u>

Es sei:

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 2 \\ 1 & 8 & 5 \\ 2 & 5 & 5 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

sowie die quadratische Funktion

$$f(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x^T A x}\right) = 2x_1^2 + 4x_2^2 + 2x_3^2 + x_1 x_2 + 5x_2 x_3 + 2x_3 x_1.$$

$$\longrightarrow \frac{\partial}{\partial \mathbf{x}} f(\mathbf{x}) = \begin{pmatrix} 4x_1 & + & x_2 & + & 2x_3 \\ x_1 & + & 8x_2 & + & 5x_3 \\ 2x_1 & + & 5x_2 & + & 4x_3 \end{pmatrix} = \begin{bmatrix} 4 & 1 & 2 \\ 1 & 8 & 5 \\ 2 & 5 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{Ax}.$$

4. Ableitung des Normquadrates einer Matrizenmultiplikation:

$$\frac{\partial}{\partial \mathbf{x}}\left(\|\mathbf{Bx}\|^2\right) = 2\mathbf{B}^T\mathbf{Bx}. \quad \text{(B-18)}$$

Beweis: Der abzuleitende Term kann folgendermassen erweitert werden:

$$\|\mathbf{Bx}\|^2 = (\mathbf{Bx})^T(\mathbf{Bx}) = \mathbf{x}^T(\mathbf{B}^T\mathbf{B})\mathbf{x}. \quad \text{(B-18a)}$$

Dann folgt (B-18) direkt aus dem Beweis von (B-17).

Unter Beachtung dieser 4 Regeln erhalten wir für unsere Lagrange-Funktion (B-14) folgende Lösungen :

$$\frac{\partial L}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x} - \lambda(2\mathbf{x})] = 0 \implies \mathbf{A}\mathbf{x} = \lambda\mathbf{x}. \tag{B-19}$$

$$\frac{\partial L}{\partial \lambda} = \| \mathbf{x} \|^2 - 1 \ = 0. \implies \| \mathbf{x} \|^2 = 1 \tag{B-20}$$

Dies impliziert, dass $\mathbf{x}$, λ ein Eigenpaar der Matrix $\mathbf{A}$ ist. Für irgendeine Lösung $\lambda = \lambda_i, \mathbf{x} \ = \ \mathbf{v_i}$, nimmt der Rayleigh-Quotient folgenden Wert an (zuerst $\mathbf{Ax} = \lambda\mathbf{x}$ beidseitig mit $\mathbf{v_i^T}$ multiplizieren):

$$\mathbf{v_i^T A v_i} = \mathbf{v_i^T}(\lambda_i \mathbf{v_i}) = \lambda_i \|\mathbf{v_i}\|^{\mathbf{2}} = \lambda_i \tag{B-21}$$

Deshalb ist der Eigenvektor $\mathbf{v_i}$ (paarweise mit dem zum grössten Eigenwert λ_i von $\mathbf{A}$) der globale Maximierer, der zum absoluten Maximum von λ_1 führt, einen grösseren Wert gibt es nicht! In gleicher Weise ist das Paar, gebildet aus dem Eigenvektor $\mathbf{v_n}$ mit dem kleinstem Eigenwert λ_n, der globale Minimierer mit absolutem Minimum λ_n, einen kleineren Wert gibt es nicht! Es gilt:

$$\lambda_{\max} \geq \lambda_1 \geq \lambda_n \geq \lambda_{\min} \tag{B-22}$$

Zusammengefasst gilt:

$$\text{Eigenpaar } (\mathbf{v_1}, \lambda_1) = \text{Maximierer von } \mathbf{x^T A x}$$
$$\text{Eigenpaar } (\mathbf{v_n}, \lambda_n) = \text{Minimierer von } \mathbf{x^T A x}$$

B.3 Generalisierter Rayleigh-Quotient

Weiter sei nun ein positiv definite Matrix $\mathbf{B} \in \mathbb{R}^{(\mathbb{m} \times \mathbb{m})}$ gegeben, d.h. sie hat die gleiche Grösse wie die aus (B-1) bekannte Matrix $\mathbf{A} \in \mathbb{R}^{(\mathbb{m} \times \mathbb{m})}$. Dann nennt man die folgende Grösse:

$$\mathbf{Q}_{\text{gen}}(\mathbf{x}) = \frac{\mathbf{x}^T \mathbf{A}\mathbf{x}}{\mathbf{x}^T \mathbf{B}\mathbf{x}} \tag{B-23}$$

den **generalisierten Rayleigh-Quotienten.**

Den Beweis des Maximierungsproblems im generaliseirten Fall zeigen wir anhand der Substitutionsmethode.

Der generalisierte Rayleigh-Quotient (B-23) wird dabei so umgewandelt, dass ein "einfacher" Rayleigh-Quotient der Form von (B-1) gebildet wird. $\mathbf{B}$ besitzt eine Quadratwurzel und kann deshalb zerlegt werden in

$$\mathbf{B} \ = \ \mathbf{B^{1/2}} \cdot \mathbf{B^{1/2}}, \tag{B-24}$$

weiter sei

$$\mathbf{y} \ = \ \mathbf{B^{1/2}x}, \text{ resp. } \mathbf{y^T} \ = \mathbf{x^T}(\mathbf{B^{1/2}})^{\mathbf{T}} \tag{B-25}$$

Somit wird der Nenner von (B-23) zu

$$\mathbf{x^T B x} \;=\; \mathbf{x^T B^{1/2} B^{1/2} x} \;=\; \mathbf{y^T y}. \tag{B-26}$$

Unter Berücksichtigung, dass $\mathbf{B^{1/2}} \cdot \mathbf{B^{-1/2}} = (\mathbf{B^{1/2}})^{-\mathbf{T}}(\mathbf{B^{1/2}})^{\mathbf{T}} = \mathbf{I}$, folgt für (B-23):

$$\begin{aligned}\mathbf{Q}_{\text{gen}}(\mathbf{x}) = \frac{\mathbf{x^T A x}}{\mathbf{x^T B x}} = \frac{\mathbf{x^T (B^{1/2})^T (B^{1/2})^{-T} \cdot A \cdot B^{-1/2} B^{1/2} x}}{\mathbf{x^T B^{1/2} B^{1/2} x}} &= \frac{\mathbf{y^T} \left[(\mathbf{B^{1/2}})^{-\mathbf{T}} \cdot \mathbf{A} \cdot \mathbf{B^{-1/2}} \right] \mathbf{y}}{\mathbf{y^T y}} \\ &= \frac{\mathbf{y^T \cdot D \cdot y}}{\mathbf{y^T y}}\end{aligned} \tag{B-27}$$

Damit wurde der Ausdruck (B-27) in Bezug auf die Matrix $\mathbf{D}$ in einen einfachen Rayleigh-Quotienten überführt. Der Beweis des Maximierungsproblems kann nun analog dem Resultat des einfachen Rayleigh-Quotienten des Absatzes 1.1 durchgeführt werden.

Allerdings lautet die Randbedingung nun:

$$\|\mathbf{B^{1/2} x}\| = \mathbf{x^T B x} = 1, \tag{B-28}$$

und (B-3) wird im Falle des generalisierten Rayleigh-Quotienten zu:

$$\max_{(\mathbf{x} \in \mathbb{R}^n : \|\mathbf{B^{1/2} x}\| = \mathbf{x^T B x} = 1)} \mathbf{x^T A x}. \tag{B-29}$$

Die Lagrange-Funktion wird zu:

$$\mathbf{L}(\mathbf{x}, \lambda) = \mathbf{x^T A x} - \lambda \underbrace{(\mathbf{x^T B x - 1})}_{\substack{\text{Randbeding-} \\ \text{ungen} = 1}}. \tag{B-30}$$

Ableitung und Nullsetzen von (B-30) liefert die Lösungen:

$$\frac{\partial L}{\partial \mathbf{x}} = 2\mathbf{Ax} - \lambda(2\mathbf{Bx} = 0 \;\longrightarrow\; \mathbf{Ax} = \lambda \mathbf{Bx} \tag{B-31}$$

$$\frac{\partial L}{\partial \lambda} = \|\mathbf{B^{1/2}} x\| - 1 = 0 \;\longrightarrow\; \|\mathbf{B}^{1/2}\mathbf{x}\| = 1. \tag{B-32}$$

Unter Verwendung von $\mathbf{B}^{1/2}$ folgt für (B-31):

$$\begin{aligned} &\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2})(\mathbf{B}^{-1/2}\mathbf{x}) = \lambda(\mathbf{B}^{1/2}\mathbf{x}) \\ &\overbrace{(\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2})}^{\mathbf{D}}\mathbf{y} = \lambda \mathbf{y} \\ &\mathbf{D}\mathbf{y} = \lambda \mathbf{y}. \end{aligned} \tag{B-33}$$

Das ist nun unser Eigenwertproblem für die Matrix **D**. Die ursprünglichen Eigenvektoren findet man durch die Rücktransformation :

$$\mathbf{x} = \mathbf{B}^{-1/2}\mathbf{y}. \tag{B-34}$$

Die direkte Berechnung von **x** aus der zugehörigen Eigenmatrix berechnet sich durch folgende Umformungen.

D in (B-33) ausschreiben und $\mathbf{y} = \mathbf{B}^{1/2}\mathbf{x}$ verwenden, ergibt:

$$\mathbf{B}^{-1/2}\mathbf{A}\underbrace{\mathbf{B}^{-1/2}\mathbf{B}^{1/2}}_{\mathbf{I}}\mathbf{x} = \lambda\mathbf{B}^{1/2}\mathbf{x}$$

$$\text{resp. } \mathbf{B}^{-1/2}\mathbf{A}\mathbf{x} = \lambda\mathbf{B}^{1/2}\mathbf{x} \tag{B-35}$$

Beide Seiten nun linksseitig mit $\mathbf{B}^{-1/2}$ multiplizieren:

$$\underbrace{\mathbf{B}^{-1/2}\mathbf{B}^{-1/2}}_{\mathbf{B}^{-1}}\mathbf{A}\mathbf{x} \;=\; \lambda\,\underbrace{\mathbf{B}^{-1/2}\mathbf{B}^{1/2}}_{\mathbf{I}}\mathbf{x}. \tag{B-36}$$

Oft wird **x** als Eigenvektor zum grössten Eigenwert λ_1 als $\mathbf{v_1}$ bezeichnet. Dann wird (B-36) zu:

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{v_1} \;=\; \lambda_1\mathbf{v_1} \tag{B-37}$$

Das Paar (λ_1, $\mathbf{v_1}$) berechnen sich als grösster Eigenwert und grösster Eigenvektor der Matrize $\mathbf{B}^{-1}\mathbf{A}$:

Da sich **x** und **y** auf gleichen Rayleigh-Quotienten beziehen, haben die Matrizen $\mathbf{B}^{-1/2}\cdot\mathbf{A}\cdot\mathbf{B}^{-1/2}$ und $\mathbf{B}^{-1}\mathbf{A}$ gleiche Eigenwerte, aber <u>verschiedene Eigenvektoren</u>.

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{v_1} \;=\; \lambda_1\mathbf{v_1} \;\Longleftrightarrow\; \mathbf{A}\mathbf{v_1} \;=\; \lambda\mathbf{B}\mathbf{v_1} \tag{B-38}$$

$$\boxed{\mathbf{A}\mathbf{v} \;=\; \lambda\mathbf{B}\mathbf{v} \text{ heisst generalisiertes Eigenwertproblem}}$$