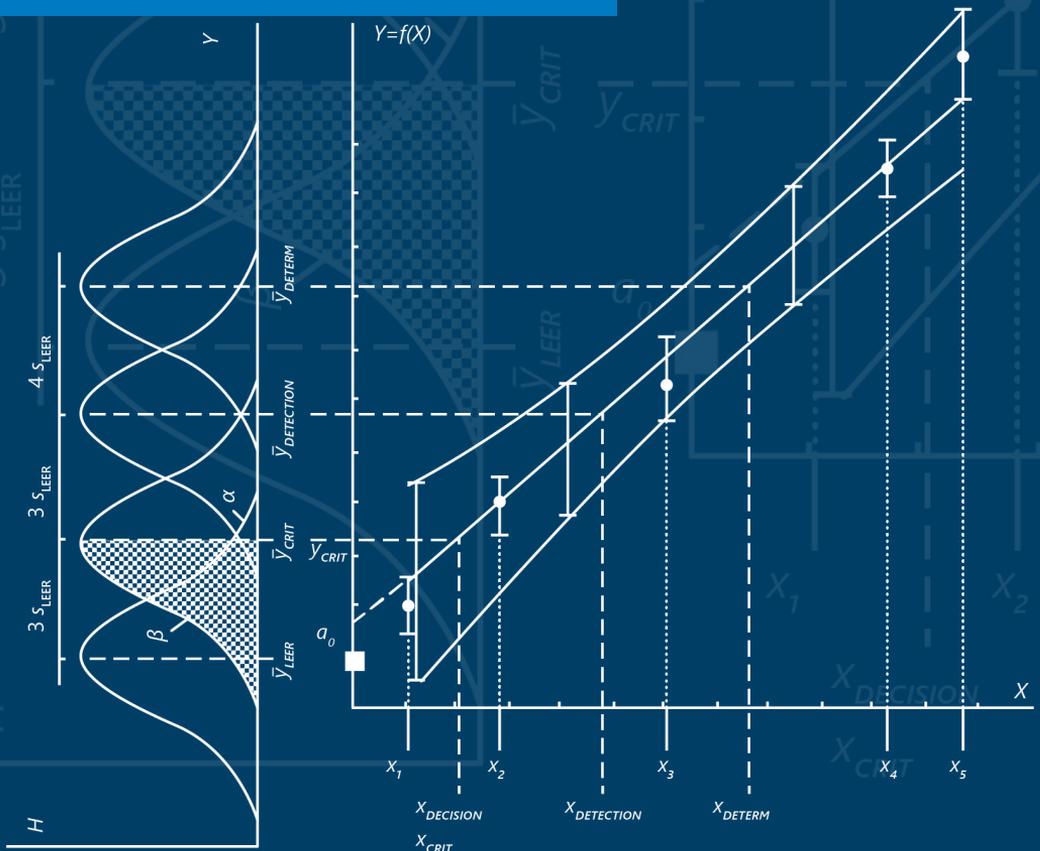


Heinz W. Zwanziger, Eduard Sorkau

# Kalibration analytischer Methoden

Theorie & Techniken

Eine Orientierungshilfe



## Impressum

### Information zur Reihe

Die Fachbuchreihe „Ingenieur- und Naturwissenschaften“ erscheint im Hochschulverlag Merseburg.  
<https://www.hs-merseburg.de/hochschule/einrichtungen/hochschulbibliothek/hochschulverlag/>

Kontakt zum Herausgeber:

[Regionalgeschichte@hs-merseburg.de](mailto:Regionalgeschichte@hs-merseburg.de)

Hochschule Merseburg, Hochschulbibliothek  
Eberhard-Leibniz-Str. 2, 06217 Merseburg

### Information zur Publikation

2021 Heinz W. Zwanziger

Autor: Heinz W. Zwanziger, Eduard Sorkau

Titel: Kalibration analytischer Methoden. Theorie & Techniken. Eine Orientierungshilfe.

DOI: <http://dx.doi.org/10.25673/38489>

163 Seiten, 22 Abb., zahlreiche Tab.

Bildrechte: Heinz W. Zwanziger

Layout: Yann Labry

Dieses Werk ist ein eBook im PDF-Format und eine freie online-Ausgabe.  
Das Werk darf heruntergeladen und als Privatkopie ausgedruckt werden.  
Es ist nicht gestattet, Kopien oder gedruckte Fassungen kommerziell zu nutzen.

Werden Auszüge aus dem Werk verwendet, ist es wie folgt zu zitieren.

Zwanziger, Heinz W., Sorkau, Eduard (2020): Kalibration analytischer Methoden. Theorie & Techniken. Eine Orientierungshilfe. In: Fachbuchreihe Ingenieur- und Naturwissenschaften. Hochschulverlag Merseburg.

DOI: <http://dx.doi.org/10.25673/38489>

## Zusammenfassung

Das Buch greift ein aktuelles Thema auf und bietet eine Einführung in die Kalibration analytischer Methoden. Ohne kalibrierte Methoden erhält man keine zuverlässigen experimentellen Daten, die man von der betrieblichen Prozessführung bis zu politischen Entscheidungen in Umweltfragen benötigt.

Die Autoren klären zunächst begriffliche Grundlagen, danach beschreiben sie gängige und alternative Kalibriermodelle sowie deren statistische Grundlagen, Bewertungen und Vergleiche, Nachweisgrenzen und spezielle Kalibriertechniken. In einem weiteren Abschnitt wird auf die Mehrkomponentenanalytik eingegangen und der Nutzen von multivariaten Auswertetechniken wie der Cluster-, Faktor- und Diskriminanzanalyse dargestellt. Anhand durchgerechneter Datenbeispiele werden die behandelten Methoden ausführlich illustriert.

Im Anhang wird eine Orientierungshilfe zur Methodenauswahl für die Multivariate Datenauswertung und die Modellierung angeboten.

Zum Verständnis dieses Buches sind keine speziellen statistischen Vorkenntnisse erforderlich.

## Stichwörter

Analytik, Einkomponentenanalytik, Mehrkomponentenanalytik, Kalibriertechniken, Kalibration, Kalibrieren, Kalibriermodelle, Modellierung, Modellbewertung, Regressionsmodelle, statistische Bewertung, Validierung, Qualitätssicherung, Fehler, Methodenauswahl, Nachweisgrenzen, graphische Auswertung, multivariate Methoden, Matrixalgebra, Diskriminanzanalyse, Signal/Rauschen-Modellierung, Mustererkennung, explorative Datenanalyse, Rechenbeispiele, Orientierungshilfe für multivariate Datenauswertung und Modellierung

## Summary

The book addresses a current topic and offers an introduction to the calibration of analytical methods. Without calibrated methods, one cannot obtain reliable experimental data which are needed throughout from operational process control to political decisions on environmental issues.

The authors first clarify conceptual basics, afterwards they describe common and alternative calibration models as well as their statistical foundations, evaluations and comparisons, detection limits, and special analytical calibration techniques. A further section discusses multi-component analysis and illustrates the utility of multivariate evaluation techniques such as cluster, factor, and discriminant analysis.

Calculated data examples are used to illustrate in detail the methods covered.

Guidance on method selection with respect to multivariate data analysis and modeling is provided in the appendix.

No special prior statistical knowledge is required to understand this book.

## Keywords

chemical analysis, analytical methods, single-component, multi-component, calibration techniques, calibration models, regression models, modeling, statistical evaluation, validation, quality assurance, errors, method selection, detection limits, graphical methods, multivariate methods, matrix algebra, discriminant analysis, signal/noise-modeling, pattern cognition, pattern recognition, exploratory data analysis, data mining, calculated examples, guidance for multivariate data evaluation and modeling

*Gewidmet Frau Dipl.-Ing. (FH) Erika Ochmann,  
Mitbegründerin des Analytik-Praktikums,*

*und Frau Dr. Barbara Wohlfarth †,  
Vertreterin der Analytik-Professur,*

*sowie unseren geduldigen Ehefrauen.*

## Zur Vorgeschichte des Buchs

---

Die erste Fassung einer Einführung in die „Kalibration analytischer Methoden“ entstand im November 1996. Für den Gebrauch im Praktikum meines Lehrgebiets wurde im Juli 1998 eine verkürzte Fassung gedruckt, im Sommersemester 2000 erschien die zweite und letzte Fassung. Bis 2000 stand Erika Ochmann zuverlässig an meiner Seite, beschaffte neue Analysentechnik und führte sie ein, testete die Arbeitsvorschriften und optimierte die Praktikumsversuche.

Danach vertrat Dr. Barbara Wohlfarth meine Professorenstelle und musste das komplette Praktikum weiter betreuen, denn im Oktober 2000 wurde ich zum Rektor der damaligen Fachhochschule Merseburg gewählt. Ganz neue Herausforderungen beanspruchten meine Zeit.

Das blieb so bis zu meiner Emeritierung zum Sommersemester 2012.

In jenem Semester versuchte ich, viel von dem wieder aufzunehmen, was ich seit über einem Jahrzehnt ruhen lassen musste. Praktische Anwendungen der Nah-Infrarot-Reflexionspektroskopie, die ohne multivariate Methoden nicht auskommt, standen zunächst im Vordergrund.

Als ich das Manuskript zum Kalibrier-Thema wieder aufnehmen wollte, standen mehrere Generationen und Wechsel von Betriebssystemen und Text-Programmen zwischen dem Urtext und einer Neufassung. Zusätzliche Probleme bereitete die wesentlich schwerfälligere Formelerstellung mit aktuellen Text-Programmen. So war das Forschungssemester eher vorbei, als ein weiteres Manuskript hätte erstellt werden können.

„Zwischendurch“ mutierte ich 2013 zum „Heimatsforscher“. Danach musste ich wegen einer schweren Krebserkrankung von 2014 bis 2017 pausieren. Schließlich fand ich aber zu neuer Schaffenskraft zurück und besann mich auf das immer noch aktuelle Kalibrier-Thema.

In dieser Annahme bestärkte mich der Amtsnachfolger in meinem Lehrgebiet an der Hochschule Merseburg, Prof. Dr. Valentin Cepus, der seit Jahren mein „Statistik-Vademekum“ für seine Lehrzwecke nutzt. Dies tat auch Dr. Eduard Sorkau, ausgewiesener Praktiker und mein erster Assistent an der Technischen Hochschule Merseburg, die von 1954-1993 existiert hatte. Im Institut für Analytik der Martin-Luther-Universität Halle hatte er ab 1993 konsequent das statistische „Handwerk“ weiter betrieben, das uns Prof. Dr. Klaus Doerffel (1925-1995) in Merseburg bis zu seinem Ruhestand lehrte und weitergab. Inzwischen hatte Eduard Sorkau den Übergang in den beruflichen Ruhestand gut bewältigt und stand für eine neue Aufgabe zur Verfügung. Hätte mir nur einer der beiden abgeraten, wäre das vorliegende Büchlein nicht entstanden.

Mit Optimismus konnte ich nun meinen „historischen“ Text gründlich überarbeiten und Eduard Sorkau als Koautor gewinnen.

Heinz W. Zwanziger

Zöschen/Leuna/Merseburg, im Januar 2020

## Vorwort zum Thema Kalibration

---

Warum glauben wir, dass die Thematik noch aktuell genug ist, um einen Text, dessen Quellenbasis aus dem Jahr 2000 stammt, wieder aufzugreifen?

In unseren Vorlesungen haben wir häufig darauf hingewiesen, dass die Analytik (Analytische Chemie und darauf beruhende Messverfahren) die wichtigste Informationsquelle ist, die wir zur materiellen Welterkenntnis haben.<sup>1</sup>

Dabei will man sich doch nicht irren? Oder sich durch eine ungesicherte Methodik der Datengewinnung auf Fehlinformation ein- und verlassen.

Deshalb muss das Instrumentarium unseres Erkenntnisgewinns brauchbar sein und zuverlässige Ergebnisse liefern. Eben dazu muss es kalibriert sein.

Selbst eine Waage, auf der wir gelegentlich unsere Körpermasse kontrollieren, oder ein Navigationssystem, von dem wir uns gerne leiten lassen, muss kalibriert werden.

In fast jedem Herstellungsprozess fallen Kalibrationsaufgaben an. Prozessanalytik, ob in der chemischen, pharmazeutischen oder Lebensmittelindustrie, braucht kalibrierte Verfahren.

Umweltanalytische und kriminaltechnische (forensische, rechtsmedizinische) Untersuchungen sind ebenfalls auf kalibrierte Verfahren angewiesen.

Dies sind nur wenige Beispiele.

Kalibration: Eine allgegenwärtige Selbstverständlichkeit also? Ja.  
Damit rechtfertigen wir die Arbeit an diesem Buch.

Allerdings ist Folgendes anzumerken:

(1) Literatur, die spezielle Aspekte des Kalibrierens erstmals behandelte, haben wir in Fußnoten gewürdigt. Buch-Publikationen oder Tabellenwerke, denen wir bleibenden Wert zuschreiben, sind in Abschnitt 7 aufgeführt.

(2) Da sich fachliche Normenwerke ständig ändern, müssen wir es dem Interessierten zumuten, sich über die aktuellen Fassungen selbst zu informieren. Auch werden die Hersteller und Vertreiber moderner instrumenteller Technik dafür sorgen, dass ihre Produkte nebst Steuerungs- und Auswertesoftware den aktuell geltenden Normen entsprechen.

(3) Rechenbeispiele, die ursprünglich in den Text eingestreut worden waren, damit die Studierenden auf leichte Weise das Besprochene selbst und per Hand nachvollziehen konnten, sind nun in drei eigenen Abschnitten zusammengefasst.

Möge die vorliegende Einführung Studierenden wie Lehrenden hilfreich sein.

Heinz W. Zwanziger  
Zöschen, im April 2020

Eduard Sorkau  
Eckartsberga, im April 2020

---

<sup>1</sup> Siehe auch: K. H. KOCH: „Analytik - (ein) Schlüssel zur Welt. Die Dynamik einer marktbeeinflussenden Informationswissenschaft und unverzichtbaren Querschnittsdisziplin“, LABO Trend 2002, 10-13

# Inhaltsverzeichnis

<b>Zur Vorgeschichte des Buchs</b>	<b>5</b>
<b>Vorwort zum Thema Kalibration</b>	<b>6</b>
<b>A Einkomponentenanalytik</b>	<b>13</b>
<b>1 Begriffliche Grundlagen</b>	<b>13</b>
1.1 Verwirrende Begriffswelt?	13
1.2 Symbole und Abkürzungen	14
1.3 Basisgrößen, abgeleitete Größen, Einheiten	16
1.4 Warum Kalibration?	17
1.5 Univariate oder multivariate Kalibration?	18
<b>2 Kalibriermodelle</b>	<b>19</b>
2.1 Das lineare Kalibriermodell	19
2.2 Die Modellparameter (Koeffizienten)	19
2.2.1 Regressionsrechnung	19
2.2.2 Voraussetzungen der Berechnung der Koeffizienten	21
2.2.3 Ausreißer	22
2.3 Statistische Bewertung der Koeffizienten	23
2.3.1 Der Fehler der Koeffizienten	23
2.3.2 Der Test der Koeffizienten	24
2.3.3 Der Vertrauensbereich der Koeffizienten (Konfidenzintervall)	25
2.4 Alternative lineare Modellierung	25
2.4.1 Berücksichtigung von Experimentalfehlern	25
2.4.2 Inverse Kalibration	27
2.4.3 Gewichtete Regression	28
2.4.4 Robuste Regression	28
2.4.5 Fuzzy Regression	29

<b>2.5 Nichtlineare Kalibriermodelle</b>	<b>30</b>
2.5.1 Modelle 1. Art (quasilineare)	30
2.5.2 Modelle 2. Art („echt nichtlineare“)	31
<b>2.6 Beurteilung von Kalibriermodellen</b>	<b>32</b>
2.6.1 Bestimmtheitsmaß ( $B$ )	32
2.6.2 Anpassungsgüte ( $GOF$ )	32
2.6.3 Anpassungsdefekt ( $LOF$ )	33
<b>2.7 Statistische Bewertung und Vertrauensbereiche</b>	<b>34</b>
2.7.1 Varianzanalyse	34
2.7.2 Vertrauensbereich der Signalwerte in der Kalibrierphase	36
2.7.3 Vertrauensbereich der Signalwerte in der Anwendungsphase	36
2.7.4 Vertrauensbereich eines Analysenergebnisses	37
<b>2.8 Qualitätssicherung und Validieren von Kalibriermodellen</b>	<b>38</b>
2.8.1 Standardreferenzmaterialien	38
2.8.2 Validierung	39
<b>2.9 Vergleich von Kalibrier- bzw. Regressionsmodellen</b>	<b>42</b>
2.9.1 Vergleich der Regressionsparameter	42
2.9.2 Vergleich der Gütemaße und Leistungsparameter	42
2.9.3 Vergleich von Regressionsmethoden	42
<b>2.10 Datenbeispiele</b>	<b>44</b>
2.10.1 Datenbeispiel zu 2.2.1 - Kalibrierdaten (Einfachmessungen)	44
2.10.2 Datenbeispiel zu 2.2.2 - Statistische Voraussetzungen	45
2.10.3 Datenbeispiel zu 2.2.3 - Ausreißertest	49
2.10.4 Datenbeispiel zu 2.3 - Statistik der Koeffizienten	50
2.10.5 Datenbeispiel zu 2.4 - Alternative Kalibriermodelle	50
2.10.6 Datenbeispiel zu 2.5 - Nichtlineare Modelle	52
2.10.7 Datenbeispiel zu 2.6 - Beurteilung von Modellen (Einfachmessungen)	52
2.10.8 Datenbeispiel zu 2.6 - Beurteilung von Modellen (Mehrfachmessungen)	53
2.10.9 Datenbeispiel zu 2.7 - Signalbezogene Statistik (Varianzanalyse)	56
2.10.10 Datenbeispiel zu 2.7 - Analytbezogene Statistik (Vertrauensbereich)	58
2.10.11 Datenbeispiel zu 2.9 - Vergleich von Regressionsgeraden	59

<b>3 Nachweisgrenzen und andere Grenzwerte</b>	<b>61</b>
3.1 Begriffliche Probleme	61
3.2 Besonderheiten in der Spurenanalytik	62
3.3 Der kritische Signalwert und die Verfahrensgrenzen	63
3.4 Die begriffliche Konfusion in der Literatur	67
<b>4 Die Kalibrier-Techniken</b>	<b>69</b>
4.1 Terminologie und Einführung	69
4.2 Externe Kalibration	70
4.2.1 Allgemeines und Prinzip	70
4.2.2 Vereinfachung durch nur einmalige Kalibrierprobenmessung	72
4.3 Analyt-Zugabe zur Testprobe	73
4.3.1 Allgemeines und Prinzip	73
4.3.2 ... mit mehreren Analyt-Zugaben	76
4.3.3 ... mit einmaliger Analyt-Zugabe	77
4.3.4 ... mit doppelter Analyt-Zugabe	79
4.4 Zugabe der Testprobe zu einer Kalibrierprobe	79
4.5 Analyt-Entfernung aus einer Testprobe	80
4.6 Analyt-Entfernung aus einer Kalibrierprobe	81
4.7 Interne Kalibrierung	82
4.7.1 Prinzip und Vorgehen	82
4.7.2 Vereinfachungen	83
4.8 Weitere „Spezialfälle“	84
4.8.1 Mehrfachprobenahme aus dem Dampfraum (GC)	84
4.8.2 ELISA-Kalibrierung	86
4.9 Vorläufige Orientierung: Welche Technik wann anwenden?	86
<b>B Mehrkomponentenanalytik</b>	<b>89</b>
<b>5 Zweikomponenten-Kalibration</b>	<b>89</b>
5.1 Modellgleichung, Signalmodellierung	89
5.2 Modellparameter, Empfindlichkeitenmatrix	91

<b>5.3 Bewertung und Auswahl eines Kalibrationsmodells</b>	<b>92</b>
5.3.1 ... nach der Empfindlichkeit	92
5.3.2 ... nach der Selektivität (Messstellen-Kriterium)	93
5.3.3 ... nach der Spezifität (Komponenten-Kriterium)	94
5.3.4 ... mittels anderer Methoden	95
<b>5.4 Anwenden des Kalibrationsmodells</b>	<b>96</b>
<b>5.5 Kenngrößen für Fehler der Kalibration</b>	<b>97</b>
<b>5.6 Maximalfehler des Ergebnisses</b>	<b>98</b>
<b>5.7 Datenbeispiele</b>	<b>100</b>
5.7.1 Datenbeispiel - Visuelle Spektrenbewertung	100
5.7.2 Datenbeispiel zu 5.2 - Modellierung, Modellbewertung	101
5.7.3 Datenbeispiel zu 5.3 - KAISERSche Auswahlkriterien	104
5.7.4 Datenbeispiel zu 5.4 - Anwenden des Kalibriermodells	105
5.7.5 Datenbeispiel zu 5.6 - Maximalfehler; Konditionszahl	106
<b>6 Mehrkomponenten-Kalibration</b>	<b>109</b>
<b>6.1 Modellierungsstrategie</b>	<b>109</b>
6.1.1 Multivariates Standard-Kalibriermodell	109
6.1.2 Multivariates inverses Kalibriermodell	111
<b>6.2 Problemvermeidungsmöglichkeiten</b>	<b>112</b>
6.2.1 Nichtkonstanz der Kalibrierkoeffizienten	112
6.2.2 Numerische Probleme	112
6.2.3 Rauschanteil beseitigen (latente Variablen)	114
6.2.4 Visuelle Ausreißerererkennung	115
6.2.5 „Wechselwirkungen“ und „Interferenzen“	116
<b>6.3 Aufwandsoptimierung</b>	<b>116</b>
6.3.1 Kalibrierproben optimal herstellen	116
6.3.2 Messstellenbewertung und -auswahl	117
6.3.3 Systemische Signal/Rausch-Optimierung	119
<b>6.4 Datenbeispiele</b>	<b>123</b>
6.4.1 Datenbeispiel zu 6.1.1 - Direkte Kalibration	123
6.4.2 Datenbeispiel zu 6.1.2 - Inverse Kalibration	124

6.4.3	Datenbeispiel zu 6.2.2 - Linear abhängige Kalibrierdaten	124
6.4.4	Datenbeispiel zu 6.2.4 - Visuelle Ausreißerererkennung ( $\lambda$ -Paare, CA, PCA)	126
6.4.5	Datenbeispiel zu 6.3.1 - Optimale Kalibrierproben (Versuchsplanung)	132
6.4.6	Datenbeispiel zu 6.3.2 - Messstellenauswahl (FA, LDA)	134
6.4.7	Datenbeispiel zu 6.3.3 - „Systemische“ Signal/Rausch-Modellierung (LDA)	138
<b>7</b>	<b>Literatur</b>	<b>145</b>
7.1	Einkomponentenanalytik	145
7.2	Mehrkomponentenanalytik	145
	<b>Anhang: Orientierungshilfe zur Methodenauswahl für Multivariate Daten- auswertung und Modellierung</b>	<b>146</b>
	<b>Verzeichnis der Abbildungen</b>	<b>150</b>
	<b>Sachwortverzeichnis</b>	<b>151</b>
	<b>Danksagung</b>	<b>163</b>

# A Einkomponentenanalytik

## 1 Begriffliche Grundlagen

---

### 1.1 Verwirrende Begriffswelt?

„Äußerlich“ betrachtet ist unser Lebensraum vierdimensional. Wir leben in drei Raumdimensionen und können unsere Erlebnisse auf einem Zeitstrahl (Zeitpfeil) als vierter Dimension einordnen. Unsere „Erlebniswelt“ (Sinneswahrnehmung) ist bereits höherdimensional. Unsere bewussten Wahrnehmungen beruhen auf dem Sehen und Hören, dem Riechen und Schmecken sowie dem Tastsinn. Mit letzterem fühlen wir; Gefühle hingegen sind daraus abgeleitete und irgendwie dem Denken geschuldete „Ergebnisgrößen“ aus den wahrgenommenen „Signalen“.

Was meinen wir aber in der Analytik mit „Dimension“?

Ein Analysenproblem kann „dimensional“ bezüglich quantitativ zu bestimmender Species (Analyte) sein. Es kann aber auch „dimensional“ bezüglich des Messraums sein, in dem wir die Signale bzw. Signalwerte erfassen.

In Analogie zur Alltagswelt liegt es nahe, unmittelbar Erlebtem, Gemessenem eine „Dimension“ zuzuschreiben, d. h. Ergebnisgrößen oder abhängigen Variablen.

Unser Problem ist: sprachliche Begriffe werden in der Regel kontextbezogen verwendet. Wer nicht vor einer Aussage, gleichgültig ob in einem Disput oder einem der folgenden Buch-Abschnitte, den Kontext klar benennt, läuft Gefahr, missverstanden zu werden bzw. Verwirrung zu stiften.

Die folgende anfängliche Festlegung der Symbole und Abkürzungen soll ein Beitrag zu begrifflicher Klarheit sein.

Nicht immer kann so offensichtlich klar formuliert werden wie „Einkomponenten-, Zweikomponenten- und Mehrkomponenten-Kalibration“.

Im Einkomponentenfall würden wir prinzipiell nur eine Messgröße brauchen.

Eine Umstellung der Kalibrierfunktion zur Berechnung von Analysenwerten ist einfach und wird keine Probleme bereiten.

Im Zweikomponentenfall werden wir aber eine Kombination von zwei unabhängigen Kalibriermodellen brauchen, um die Konzentration zweier Analyte in einer Probe zu bestimmen.

Obwohl jedes dieser Kalibriermodelle, bezogen auf die Zahl der modellierten Signale der abhängigen Messvariable, als *univariat* zu bezeichnen wäre, wird aus dieser Kombination zweier unabhängiger univariater Modelle ein *bivariates* Modellsystem (Gleichungssystem).

Bezüglich der Zahl unabhängigen (gleichwohl natürlich messbaren) Einflussgrößen, der beiden „Analyt-Faktoren“, die unser univariates Signal hervorrufen, kann das System aber ebenfalls als *bivariat* bezeichnet werden. (Bei der Parameterschätzung entsprechender Regressionsmodelle wäre dies angemessen; hier hat sich aber die Bezeichnung *multiple* Regression eingebürgert.) *Mehrkomponentenanalytik* wäre also notwendig *multivariat*.

Deshalb entschließen wir uns, bestimmte Größen  $X$ , wie Konzentrationsgrößen als Faktoren zu bezeichnen, während eine Anzahl von  $P$  Messgrößen einen  $P$ -dimensionalen Messraum aufspannt. Im Gegensatz zur eindimensionalen Zweikomponenten-Modellierung wäre eine Zweikomponenten-Analyse folglich auch eine zweidimensionale.

Versuchen wir, uns im Folgenden daran zu halten. Auch wenn es anderswo anders gehandhabt wird.

In der „Makrowelt“ der Analytik spräche man aber dann bei der Kopplung einer HPLC-Trennung mit Registrierung der UV-Spektren der getrennten Komponenten nicht ganz korrekt von zweidimensionaler (Mehrkomponenten-)Analytik.

Beachten wir also am besten immer den Kontext.

## 1.2 Symbole und Abkürzungen

Wenn in Teilen des folgenden Textes nicht anders definiert, haben Symbole und Abkürzungen immer die gleiche Bedeutung und sollen sich möglichst selbst erklären.

(1) Mess- und andere **Größen** (engl.: features) bzw. Variablen (engl.: variates, variables), vor allem in Modellgleichungen, werden durch **Groß**buchstaben symbolisiert.

$X$  Analytgröße; Einflussgröße, unabhängige Variable (zum Beispiel Konzentration)  
 $Y$  Signalgröße; Ziel- oder Antwortgröße, abhängige Variable (zum Beispiel Absorbanz)

(2) Mess- und Rechen**werte** (engl.: values) einer Größe werden durch **Klein**buchstaben symbolisiert. Die Natur des Wertes wird im Allgemeinen durch einen Index näher beschrieben, der wegen besserer Lesbarkeit meist großgeschrieben wird.

$x$  Konzentrationswerte („Konzentrationen“)  
 $x_C, x_{CAL}$  Konzentrationswert des Analyten in der Kalibrierprobe  
 $x_{CRIT}$  kritischer Konzentrationswert  
 $y$  Signalwerte (Messwerte) (engl.: response value, response)  
 $y_{LEER}$  Leerwert (engl.: blank value)  
 $y_{CRIT}$  kritischer Signalwert  
 $\Delta$  Differenz von Werten

(3) Konstante Zahlenwerte werden durch Großbuchstaben symbolisiert.

$C$  Zahl der Komponenten (Analyte)  
 $N$  Anzahl von Objekten, Versuchen, Zahl aller Messungen im Kalibrierschritt  
 $N_A$  Zahl der Messungen im Anwendungsschritt (Parallelbestimmungen)  
 $N_L$  Zahl der Leerwertmessungen  
 $N_C, N_{CAL}$  Zahl der Kalibrierstellen  $x_C$  bzw.  $x_n$  ( $n = 1, \dots, N_C$ )  
 $N_{REP}$  Zahl der Kalibrierstellen, an denen  $Y$  wiederholt gemessen wurde ( $n = 1, \dots, N_{REP}$ ); idealerweise ist  $N_{REP} = N_C$   
 $W_{REP}$  Anzahl der Wiederholungen ( $w = 1, \dots, W_{REP}$ ) an jeder Stelle  $n \in N_{REP}$   
 $N_P$  Zahl der Parameter in (Regressions-)Gleichungen  
 $P$  Anzahl von Messgrößen oder Messstellen einer Größe (zum Beispiel Wellenlängen) kontextbezogen auch Zahl der Koeffizienten („Parameter“) in Gleichungen

(4) Variable Zahlenwerte, Laufindizes, zum Beispiel in Summen, schreiben wir klein.

$c$  bezogen auf  $N_c$   
 $i, j, k, l, m, n$  allgemeine Laufindizes

(5) Für statistische Größen bzw. Parameter (Rechenwerte) werden meist die üblichen Symbole oder Kürzel verwendet.

$\alpha$  (Irrtums-)Wahrscheinlichkeit (des Risikos einer Fehlentscheidung erster Art)  
 $\beta$  (Irrtums-)Wahrscheinlichkeit (des Risikos zweiter Art)  
 $\mu$  Mittelwert der Grundgesamtheit, zum Beispiel unendlich vieler Messwerte von  $X$   
 $\nu$  Freiheitsgrad (Symbol nach ISO)<sup>2</sup>  
 $\sigma$  Standardabweichung einer Größe  
 $\hat{\mu} \hat{\sigma}$  (hypothetische) Schätzwerte für  $\mu$  und  $\sigma$   
 $\bar{x} \bar{y} \bar{z}$  (berechnete) Mittelwerte  
 $\hat{a} \hat{b} \hat{c} \hat{u} \hat{x} \hat{y} \hat{z}$  (berechnete) Schätz- oder Vorhersagewerte  
 $F(df_1, df_2; q)$  Werte (Quantile) der  $F$ -Verteilung (FISHERSche Verteilung)  
 $fg \ df$  Freiheitsgrad(e), meist eine Funktion von  $N$  (engl.: degrees of freedom)  
 $k$  „statistischer Faktor“, Quantil; meist eine Funktion von  $P$  oder  $P$  und  $N$   
Man spricht auch von „Fraktile der Ordnung  $q$  und  $df$  Freiheitsgraden“.  
 $P, p$  Wahrscheinlichkeit (engl.: probability), Wahrscheinlichkeitswerte  
 $q$  Ordnung eines Quantils; eine Funktion von  $\alpha$   
 $s$  (berechnete) Standardabweichung (engl.: sample standard deviation)  
 $s^2$  (berechnete) empirische Varianz (engl.: sample variance)  
 $t(df; q)$  Werte (Quantile) der  $t$ -Verteilung (STUDENTsche Verteilung)  
Nach ISO 3534-1977 (ISO Standards Handbook 3) sind die Quantile korrekt zu schreiben als:  $t_q(df)$ .  
 $u(q)$  Werte (Quantile) der Normalverteilung (GAUSSsche Verteilung)

(6) Im günstigsten Fall kann man an Kürzeln schon erkennen, was sie bedeuten („mnemotechnische“ Kürzel). Sie werden bei Verwendung in Indices mit Großbuchstaben geschrieben.

$cnf$  Konfidenzbereich (Vertrauensbereich (engl.: confidence region))  
 $PRED$  Vorhersage- oder Schätzwert (engl.: prediction)  
 $RES$  Residuum (Differenz zwischen einem berechneten und gemessenem Wert)

(7) Symbole in linearen Kalibrier- und Regressionsmodellen

$a_0$  oder  $b_0$  Achsenabschnitt (engl.: intercept, offset), „Blind-“ oder Leerwert  
 $a_1$  oder  $b_1$  Anstieg oder Empfindlichkeit (engl.: slope, sensitivity)  
 $MAQ$  mittlere Abweichungsquadrate (Quotienten aus  $SS$ -Werten und Freiheitsgraden)  
 $SAQ, SS$  Summe von Abweichungsquadraten (engl.: sum of squares)  
 $SS_{XX}$  Quadratesummen der zentrierten Messwerte von  $X$   
 $SS_{YY}$  Quadratesummen der zentrierten Messwerte von  $Y$   
 $SS_{XY}$  Summen der Produkte zentrierter  $X$ -Werte und zentrierter  $Y$ -Werte

---

2 ISO: International Organization for Standardization

### 1.3 Basisgrößen, abgeleitete Größen, Einheiten

**Basisgrößen**<sup>3</sup> nach dem internationalen Einheitensystem (SI) und ihre **Einheiten** sind:

Basisgröße	Größenzeichen	Einheit nach SI	Einheitenzeichen
Masse	$m$	Gramm	g
Volumen	$V$	Kubikmeter	$m^3$ bzw. $dm^3$
Stoffmenge	$n$	Mol	mol

In der Analytik ist eine davon abgeleitete Größenart besonders wichtig, der **Gehalt**:

Gehaltsgröße	Größenzeichen	Einheit	Definitionsgleichung
Massengehalt (Massenanteil)	$w$	$[w] = 1$	$w_A = m_A / m$
Volumengehalt (Volumenanteil)	$\varphi$	$[\varphi] = 1$	$\varphi_A = V_A / V$
Stoffmengengehalt (Molenbruch)	$x$	$[x] = 1$	$x_A = n_A / n$

Bezogen wird auf „Summe der gleichartigen Größen für alle Stoffe der Mischphase“.

Die Gehaltsangabe kann erfolgen

- durch Dezimalbruch: zum Beispiel  $w_A = 0,001$
- durch Verhältnisbezeichnung: zum Beispiel  $w_A = 0,1 \% = 1000 \text{ ppm}$
- durch Quotient gleichartiger Einheiten: zum Beispiel  $w_A = 1 \text{ mg g}^{-1} = 1 \text{ g kg}^{-1}$

Eine weitere abgeleitete Größenart ist die *Konzentration*. Sie ist nicht nur in der Konzentrationsanalytik gebräuchlich.

Konzentrationsgröße	Größenzeichen	Einheitenzeichen	Definitionsgleichung
Massenkonzentration	$\beta$	$g \text{ dm}^{-3}, \text{ mg dm}^{-3}, \dots$	$\beta_A = m_A / V$
Volumenkonzentration	$\sigma$	Vol%	$\sigma_A = 100 V_A / V$
Stoffmengenkonzentration	$c$	$\text{mol l}^{-1}$	$c_A = n_A / V$

3 Nach dem internationalen Einheitensystem (SI) bzw. in der Physikochemie/Thermodynamik ist ggf. daran zu erinnern:

- *extensive Größen* sind additiv (zum Beispiel Masse, Volumen)
- *intensive Größen* sind nicht additiv (zum Beispiel Dichte, Konzentration)
  - *stoffeigene Größen* (Dichte)
  - *systemeigene Größen* (Temperatur T, Druck p)
- *masseabhängige Größen*:
  - bei Bezug auf beliebige Masse: Kleinbuchstaben verwenden
  - wenn auf 1 Mol bezogen: Großbuchstaben verwenden

## 1.4 Warum Kalibration?

In der Analytik gibt es nur wenige *Absolutmethoden* zur Massebestimmung von Analyten (Gravimetrie, Coulometrie) in einem Medium, einer physikalischen Matrix. Solche Methoden beruhen auf der Messung von *Basisgrößen* wie Masse und Zeit.

Obwohl hierbei die Messmittel, zum Beispiel die Waagen, auch kalibriert („geeicht“) werden müssen, wird die jeweilige analytische Methode als Absolutmethode bezeichnet.

Bei allen anderen Analysenmethoden ist man aufs *Kalibrieren* (*Kalibration*) angewiesen. Wir wollen unter Kalibration einen Prozess verstehen, in dem Messwerte methodentypischer Messgrößen  $Y$  (Signalgrößen) in Abhängigkeit von einer Analytgröße  $X$ , zum Beispiel dem Gehalt oder einer entsprechenden proportionalen Größe in festen, flüssigen oder gasförmigen Messproben erfasst werden.

Diese Abhängigkeit der Signalgröße von der Analytgröße wird zunächst modelliert und anschließend validiert, d. h. auf ihre statistische Signifikanz und praktische Verwendbarkeit geprüft. Dabei setzen wir, in unserem Fachgebiet jedenfalls, voraus oder nehmen an, dass wir durch das Modell tatsächlich eine *Kausalbeziehung* beschreiben. Das muss nicht immer zutreffen: Auch ein perfekt passendes Modell muss nicht notwendig „wahr“ sein. (Mitunter wird darauf verwiesen, dass man eine wirkliche „Ursachenmodellierung“ nur mit *Strukturgleichungen* erreicht.)

Man vergleiche dazu das ANSCOMBE-Quartett in Abschnitt 2.10.2.

Generell sollten die Kalibrierproben den Analyten möglichst in einem Medium (Matrix) enthalten, dessen restliche Zusammensetzung den zu erwartenden Analysenproben weitgehend entspricht.

Steht nur ein einziges „*Standardmaterial*“ bzw. eine „*Standardprobe*“ zur Verfügung, so sind die Kalibrierproben eher durch *Aufstocken* einer Leerprobe als durch *Verdünnen* des Standards herzustellen. Bei Mehrkomponenten-Problemen sollten statt Verdünnungsreihen besser wohldurchdachte Mischungen der Analyte verwendet werden.

Wir gehen hier von Kalibrierproben aus, bei denen *Probenahme- und Probevorbereitungsfehler*<sup>4</sup> vernachlässigbar sind.

Unter *Kalibrierproben* und *Validierproben* verstehen wir solche Messproben, die sich im Wert der Analytgröße unterscheiden und von denen der Wert der Analytgröße bekannt ist. Diese Messproben liefern uns die Messwerte.

Unter *Messwerten* bzw. Messergebnissen verstehen wir eine Menge  $\{y_1, y_2, \dots, y_N\}$  von  $N$  gemessenen Werten, kurz  $y_n$  ( $n = 1, 2, \dots, N$ ), einer Signalgröße  $Y$  (zum Beispiel Extinktion bzw. Absorbanz in der Photometrie), die aus vorgegebenen Werten einer Analytgröße  $X$  (zum Beispiel Konzentration) resultieren.

Aus den Messwerten der Kalibrierproben werden dann die Parameter einer passenden *Kalibrierfunktion*  $Y = f(X)$  ermittelt (statistisch: geschätzt). Man nennt diese Funktion auch Kalibriermodell.

---

4 Zur Nomenklatur der Probenahmeschritte und der aus den Probevorbereitungsschritten entstehenden Teilmengen siehe W. HORWITZ: „Nomenclature for sampling in analytical chemistry“, Pure & Appl. Chem. 62(1990)1193-1208.

Die Umkehrfunktion der Kalibrierfunktion, die *Analysenfunktion*  $X = f(Y)$ , wird zum Berechnen von Konzentrationswerten  $x_A$  in der Anwendungsphase benötigt, also auf Validierproben und Analysenproben gleichermaßen angewendet.

## 1.5 Univariate oder multivariate Kalibration?

Im Teil A besprechen wir die *univariate Kalibration*. Ihr liegt die Auswertung von Signalen *nur einer Messstelle* zugrunde. Bezüglich der Messgrößen, besser Messstellen, handelt es sich zugleich um eine „univariate Kalibration“.

Um mehrere Komponenten gleichzeitig bestimmen zu können, muss man multivariat kalibrieren. Zur simultanen Mehrkomponenten-Kalibration (SMK) bedarf es also der (quasi) gleichzeitigen Auswertung mehrerer Signalgrößen bzw. der Messung an mehreren Messstellen einer Messgröße. Dies wird Gegenstand von Teil B sein.

Eine multivariate Methode muss aber nicht notwendig zur simultanen Mehrkomponentenanalytik verwendet werden. Insbesondere die *faktoriellen Methoden* erlauben es, die einzelnen Komponenten (Analyte) unabhängig voneinander zu bestimmen, also eine sequentielle Mehrkomponenten-Analytik zu betreiben.

Tatsächlich werden multivariate Analyse-Methoden häufig zur Einkomponenten-Bestimmung benutzt.

Zum Einstieg bzw. ergänzenden Lektüre seien die Übersichten von MARTENS und NAES (1984), NAES und MARTENS (1984), BROWN und OBREMSKI (1984), BEEBE und KOWALSKI (1987) sowie Kapitel 4 in SHARAF et al. (1986) und die durchgerechneten Beispiele in einigen Kapiteln in MASSART et al. (1988) empfohlen. Zur Vertiefung kann das Buch von MARTENS und NAES (1989) dienen. Die vollständigen Literaturzitate findet man in Abschnitt 7.

## 2 Kalibriermodelle

---

### 2.1 Das lineare Kalibriermodell

Ziel des Kalibrierens ist also das Aufstellen und statistische Absichern des mathematischen Modells der funktionalen Abhängigkeit der Signalgröße  $Y$  von der Analytgröße  $X$ .

❶ Abbildung im Datenbeispiel Abschnitt 2.10.1

Am weitesten verbreitet sind lineare Modelle des Typs

$$Y = a_0 + a_1 X \quad (1)$$

Unter Kalibrieren im engeren Sinne versteht man die Ermittlung der beiden Modellparameter (Regressionskoeffizienten), insbesondere des *Anstiegs*  $a_1$  der Geradengleichung (1), der in der Analytik auch als *Empfindlichkeit* bezeichnet wird.

### 2.2 Die Modellparameter (Koeffizienten)

#### 2.2.1 REGRESSIONSRECHNUNG

In der Regressionsanalyse bezeichnet man  $X$  als unabhängige Einflussgröße, Prädiktorvariable oder *Regressor* und  $Y$  als abhängige Antwortgröße oder *Regressand*.

Üblicherweise werden die beiden Koeffizienten in Gleichung (1) durch Regressionsrechnung beschafft, wozu zahlreiche spezielle oder in einer Tabellenkalkulationssoftware integrierte Rechenprogramme existieren. Bei wenigen Daten kann die Ermittlung auch relativ schnell per Hand erfolgen.

Wir gehen zunächst davon aus, dass zu jeder der  $N_c$  Kalibrierproben, zum Beispiel zu jeder Konzentrationsstufe, nur eine Signalmessung vorliegt.

Damit liegen  $N_c$  Wertepaare  $(x_n, y_n)$  ( $n = 1, 2, \dots, N_c$ ) an den Kalibrierstellen vor, aus denen man das Wertepaar der arithmetischen Mittelwerte  $(\bar{x}, \bar{y})$  berechnen kann. Bei Mehrfachmessungen an einer oder jeder Kalibrierstelle erhält man außerdem Mittelwerte  $\bar{y}_n$ .

❶ Abbildung im Datenbeispiel Abschnitt 2.10.8

Die Modellgleichung (1) sollte im Idealfall für jedes dieser Wertepaare gelten, d. h.

$$y_n = a_0 + a_1 x_n \quad (2)$$

In der Realität liegt natürlich nicht jeder Messpunkt auf der Kalibriergeraden.

Es sei hier auch darauf hingewiesen, dass man Größen- und Wertegleichung nicht verwechseln sollte: In Größengleichungen bedeuten die Formelzeichen Größen, in (Zahlen-) Wertegleichungen bedeuten die Formelzeichen Zahlenwerte. Wir werden versuchen, den Unterschied durch Groß- bzw. Kleinschreibung der Formelzeichen (Symbole) zu verdeutlichen.

Mit den Summen der quadratischen Abweichungen

$$SS_{XX} = \sum_{n=1}^{N_C} (x_n - \bar{x})^2 \quad (3)$$

$$SS_{YY} = \sum_{n=1}^{N_C} (y_n - \bar{y})^2 \quad (4)$$

bzw. der Summe der Abweichungsprodukte

$$SS_{XY} = \sum_{n=1}^{N_C} (x_n - \bar{x})(y_n - \bar{y}) \quad (5)$$

kann sehr einfach der Regressionskoeffizient  $a_1$  (Anstieg, Empfindlichkeit, engl.: slope) berechnet werden:

$$a_1 = \frac{SS_{XY}}{SS_{XX}} \quad (6)$$

Da alle linearen Kalibriergeraden durch den Datenschwerpunkt  $(\bar{x}, \bar{y})$  verlaufen, erhält man aus der Modellgleichung (1) durch Einsetzen der Mittelwerte und Umformen sofort die Konstante  $a_0$  (Achsenabschnitt, engl.: offset; intercept, blank value).

Genaugenommen sind  $a_0$  und  $a_1$  Schätzwerte der wahren, aber unbekanntenen Kenngrößen des Kalibriermodells, das seinerseits für einen *Kausalzusammenhang* zwischen  $Y$  und  $X$  steht. Korrekt müsste man also  $\hat{a}_0$  und  $\hat{a}_1$  schreiben.

❶ Datenbeispiel siehe Abschnitt 2.10.1

Die nach Modell nach Gleichung (1) für alle vorgegebenen  $x_n$  berechneten  $\hat{y}_n$

$$\hat{y}_n = \hat{a}_0 + \hat{a}_1 x_n \quad (7)$$

werden im allgemeinen von den gemessenen  $y_n$  abweichen:

$$y_{RES, n} = y_n - \hat{y}_n \quad (8)$$

Der einfachste Algorithmus zur Schätzung der optimalen Koeffizienten, die Methode der kleinsten Quadrate (engl.: OLS, ordinary least squares), beruht auf der Optimierung der Summe der Quadrate eben dieser Abweichungen (engl.: residual sum of squares):

$$RSS = \sum_{n=1}^{N_C} y_{RES, n}^2 \Rightarrow \min \quad (9)$$

❶ Datenbeispiel siehe Abschnitt 2.10.1

## 2.2.2 VORAUSSETZUNGEN DER BERECHNUNG DER KOFFIZIENTEN

Die korrekte Ermittlung (numerisch zuverlässige Berechnung) der Koeffizienten im Modell nach Gleichung (1) beruht auf einigen Annahmen bzw. Bedingungen:

- (1) Die Größe  $X$  lässt sich an den Stellen  $x_n$  fehlerfrei einstellen.  
Vor allem bei spurenanalytischen Methoden muss mit einer Nichterfüllung dieser Bedingung gerechnet werden.

**Ausweg:** die Kenngrößen  $a_0$  und  $a_1$  im Rahmen eines alternativen linearen Modells berechnen

- (2) Die Größe  $Y$  ist (an den Stellen  $x_n$ ) normalverteilt.

**Prüfung:** qualitativ: Abschätzung evtl. durch einfaches Auftragen von  $y_n$  gegen  $n$ ; quantitativ: durch *Anpassungstest* oder Q-Q-Plot

**Ausweg:** evtl. Transformation der (Signal-)Daten  
Anschließend sollte man unbedingt prüfen, ob (3) erfüllt ist.

- ❶ Abbildung (Q-Q-Plot) im Datenbeispiel Abschnitt 2.10.2

- (3) Die Varianz von  $Y$  (an den Stellen  $x_n$ ) ist im Kalibrierbereich konstant (homoskedastisch).

**Prüfung:** mittels COCHRAN-Test<sup>5</sup>

**Ausweg:** zum Beispiel auf *gewichtete Regressionsmodelle* ausweichen

- ❶ Abbildung im Datenbeispiel Abschnitt 2.10.2

- (4) Die Größe  $Y$  ist an den Stellen  $x_n$  ausreißerfrei.

**Prüfung:** mit dem DIXON-Test<sup>6</sup>

**Ausweg:** zum Beispiel eine *robuste* Regressionsmethode benutzen

- ❶ Abbildung im Datenbeispiel Abschnitt 2.10.2

- (5) Die Residuen  $y_{\text{RES}, n}$  sind normalverteilt und korrelieren nicht mit den  $x_n$ .

**Prüfung:** qualitativ: durch einfaches Auftragen von  $y_{\text{RES}, n}$  gegen  $x_n$   
Die Residuen sollten zufällig um Null streuen.

quantitativ: durch Anpassungstest oder Q-Q-Plot

Mittels DURBIN-WATSON-Test kann geprüft werden, ob die Residuen autokorreliert sind. (vgl. zum Beispiel MAGER, 1982)

**Ausweg:** zum Beispiel *nichtlineare* Analyt-Signal-Beziehung modellieren

- ❶ Abbildung im Datenbeispiel Abschnitt 2.10.2

---

5 DIN ISO 5725 (April 1988): „Präzision von Messverfahren; Ermittlung der Wiederhol- und Vergleichspräzision von festgelegten Messverfahren durch Ringversuche“, siehe ISO Standards Handbook 3, S. 427.

6 Ebd., S. 459.

- (6) Natürlich wird überhaupt die Berechtigung eines linearen Modells des Typs Gleichung (1) angenommen.

**Prüfung:** Tests (auf Bestimmtheitsmaße, GOF-, LOF-Tests; siehe unten)

**Ausweg:** *nichtlineare* Kalibrierung, im einfachsten Fall mit quasilinearen Modellen, also Berücksichtigung höherer Potenzen von  $X$  (Polynome) oder auch durch stückweise lineare Kalibration<sup>7</sup>

❶ Abbildung im Datenbeispiel Abschnitt 2.10.2

- (7) Gelegentlich findet man den Hinweis, dass die Zahl der Proben (Fälle) mindestens das 10- bis 20-fache der Zahl der Prädiktorvariablen betragen sollte, um stabile Schätzwerte für die Regressionsparameter zu erhalten. Auf die in diesem Zusammenhang möglicherweise auftretende Überanpassung hat LIEBSCHER<sup>8</sup> hingewiesen. (Zum Einfluss des Stichprobenumfangs auf statistische Entscheidungen siehe ZWANZIGER et al. (1986)<sup>9</sup>, zur Probenanzahl vgl. Abschnitt 6.3.1)

Man sollte sich der genannten Bedingungen immer bewusst sein. Manchmal kann die unentdeckte Nichteinhaltung einer Voraussetzung zu einem folgenschweren Mangel an Vorhersagegüte einer Kalibrierbeziehung führen.

Außerdem ist daran zu erinnern, dass die Voraussetzungen (2) bis (5) nur geprüft werden können, wenn an jeder Kalibrierstelle Wiederholungsmessungen vorliegen.

Auch wenn alle genannten Bedingungen zutreffen, kann es ratsam sein, auf eine explizite Modellierung der Kalibrierbeziehung für einen sehr weiten Wertebereich zu verzichten und, meist unter Annahme wenigstens stückweiser linearer Abhängigkeiten, spezielle Kalibriertechniken anzuwenden (siehe auch Abschnitt 4).

Es ist vielleicht interessant, darauf hinzuweisen, dass es spezielle Regressionsmodelle, *Modelle II* gibt, bei denen angenommen bzw. vorausgesetzt wird, dass  $X$  und  $Y$  normalverteilt und fehlerfrei sind.

❶ Datenbeispiel siehe Abschnitt 2.10.2

### 2.2.3 AUSREIßER

Die Ausreißerfreiheit hatten wir im vorigen Abschnitt als Voraussetzung der zuverlässigen Berechnung der Koeffizienten erwähnt.

Wenn ein Messwert in einer Gruppe zusammengehörender  $N$  Messwerte durch systematische Einflüsse verfälscht wird, dann bezeichnet man ihn als Ausreißer. Nichterkannte Ausreißer beeinflussen die Präzision einer Datenreihe. Um nicht dem Bauchgefühl nach-

<sup>7</sup> L. M. SCHWARTZ: „Nonlinear Calibration“, Anal. Chem. 49(1977)2062-2068.

<sup>8</sup> E. LIEBSCHER: „Besitzt man weniger als etwa 20 Datenwerte je Parameter, dann kommt man in den Bereich des „Overfitting“ (Überanpassung). Oft wird dies durch die Signifikanztests der einzelnen Regressionsparameter erkannt.“, pers. Mitteilung, 25.03.2020.

<sup>9</sup> H. ZWANZIGER, U. ROHLAND, G. WERNER: „Risiko 2. Art und Stichprobenumfang bei statistischen Tests“, Fresenius Z. Anal. Chem. 323(1986)371-374.

zugeben und „unliebsame“ Messwerte einfach zu eliminieren, wird ein *Ausreißertest* herangezogen.

Der DIXON-Test wird von der DIN 53804<sup>10</sup> empfohlen für  $N < 30$ . Zunächst sind die Messwerte nach steigenden Werten zu ordnen:  $y_1 \leq y_2 \leq \dots \leq y_{N-1} \leq y_N$ . Nur der kleinste Messwert,  $y_1$ , oder der größte Messwert,  $y_N$ , kommen als Ausreißer in Frage. Dann bildet man für beide Messwerte den Quotienten aus „Verdachtswert minus Nachbarwert“ und der Spannweite. Die Position des Nachbarwerts hängt von  $N$  ab. Der größte dieser beiden Quotienten wird gegen einen kritischen Höchstwert  $d_{\text{CRIT}}$  getestet und wie üblich wird über die Hypothese der Ausreißerfreiheit entschieden. Die kritischen Höchstwerte und die Berechnung der Quotienten entnimmt man geeigneten Tabellen unter Berücksichtigung des Stichprobenumfanges. In Abschnitt 2.10.3 geben wir die Tabelle nach DEAN und DIXON (1951) an.

❶ Datenbeispiel siehe Abschnitt 2.10.3

## 2.3 Statistische Bewertung der Koeffizienten

### 2.3.1 DER FEHLER DER KOEFFIZIENTEN

Die erhaltenen Geradenkenngrößen müssen auf ihre statistische Signifikanz geprüft werden und sollten mit einem Vertrauensintervall angegeben werden. Ihre Standardabweichung ist eine Funktion der Reststandardabweichung  $s_{\text{RES}}$  des Regressionsmodells. Diese ist für das Modell nach Gleichung (1) definiert als:

$$s_{\text{RES}} = \sqrt{\frac{\sum_{n=1}^{N_C} (y_n - \hat{y}_n)^2}{N_C - 2}} \quad (10)$$

Der Nenner (Freiheitsgrad) ist abhängig von der Zahl der geschätzten Regressionskoeffizienten  $N_p$  und lautet allgemein  $N_C - N_p$ .

Man beachte dies im Falle eines *homogenen Regressionsmodells*, in dem der Absolutterm in Gleichung (1) fehlt. Das kann sachlich oder statistisch begründet sein; letzteres nämlich dann, wenn  $a_0$  nicht signifikant von Null verschieden ist.

10 nach DIN 53804, Teil 1: „Statistische Auswertungen; Messbare (kontinuierliche) Merkmale“, Beuth Verlag, Berlin, Köln. Kritische Höchstwerte findet man in ISO 5725-1986 (engl. Version), in DIN 5725-1988 (dt. Version) oder bis  $N = 40$  in CAULCUTT und BODDY (1983, S. 248). MILLER und MILLER (1994, S. 224) geben Werte bis  $N = 10$  an, die aber nur bis  $N = 7$  näherungsweise mit den DIN-Werten übereinstimmen. DIN ISO 5725 präzisiert, dass es sich bei diesem DIXON-Test um die Version von R. S. GARDNER handelt.

Der Fehler, genauer die *Standardabweichung des Koeffizienten*  $a_0$  ist als Spezialfall unmittelbar aus der Standardabweichung der Vorhersagewerte  $\hat{y}_n$ , d. h. der Punkte auf der Kalibriergeraden,

$$s(\hat{y}_n) = s_{RES} \sqrt{\frac{1}{N_C} + \frac{(x_n - \bar{x}_C)^2}{SS_{XX}}} \quad (11)$$

abzuleiten:

$$s(\hat{a}_0) = s_{RES} \sqrt{\frac{1}{N_C} + \frac{\bar{x}_C^2}{SS_{XX}}} \quad (12)$$

Die *Standardabweichung des Anstiegs* ist

$$s(\hat{a}_1) = \frac{s_{RES}}{\sqrt{SS_{XX}}} \quad (13)$$

In den obigen Gleichungen ist stets  $N_C = N$ .

❶ Datenbeispiel siehe Abschnitt 2.10.4

### 2.3.2 DER TEST DER KOEFFIZIENTEN

Für den Test der Koeffizienten wird zunächst als Testgröße

$$t_{EMP} = \frac{|\hat{a}_p - 0|}{s(\hat{a}_p)} \quad (14)$$

für jeden der  $P$  Koeffizienten berechnet.

Der Wert dieser Größe wird gegen  $t_{CRIT}(N_C - N_P; 1-\alpha/2)$  getestet, und es wird wie üblich entschieden: Wenn  $t_{EMP} \geq t_{CRIT}$ , dann ist der Koeffizient signifikant von Null verschieden.

Nach MAGER (1982, S. 37) sind, wenn keine Wiederholungsmessungen vorliegen, die Tests von  $a_1$ , des Bestimmtheitsmaßes  $B$  (siehe Abschnitt 2.6.1) und der Anpassungsgüte  $GOF$  (siehe Abschnitt 2.6.2) einander völlig gleichwertig.

Ist das Testergebnis für  $a_0$  „nicht signifikant“ und das für  $a_1$  „signifikant“, so muss ein homogenes Regressionsmodell als wahrscheinlich angenommen werden:

$$Y = a_1 X \quad (15)$$

In Ringversuchen bzw. bei Methodenvergleichen kann auch ein Vergleich von  $a_1$  mit dem Wert 1 sinnvoll sein.

❶ Datenbeispiel siehe Abschnitt 2.10.4

### 2.3.3 DER VERTRAUENSBEREICH DER KOEFFIZIENTEN (KONFIDENZINTERVALL)

Für die signifikant vom Vergleichswert verschiedenen Koeffizienten lassen sich nun deren Vertrauensbereiche angeben, wobei man von Gleichung (14) ausgeht und berücksichtigt, dass:

$$t_{EMP} < t_{CRIT} \quad (16)$$

Für das Vertrauensintervall ergibt sich die folgende Beziehung, wobei  $a_{WAHR}$  für den wahren aber unbekanntem Koeffizienten steht und  $k$  die entsprechenden Quantile sind:

$$a_p - k s(a_p) < a_{WAHR} < a_p + k s(a_p) \quad (17)$$

❶ Datenbeispiel siehe Abschnitt 2.10.4

## 2.4 Alternative lineare Modellierung

### 2.4.1 BERÜCKSICHTIGUNG VON EXPERIMENTALFEHLERN

Wenn damit zu rechnen ist, dass beide Größen in Gleichung (1) fehlerbehaftet sind, muss man nach alternativen Berechnungsmöglichkeiten der Kalibrierkoeffizienten suchen. Das Verhältnis der „wahren Experimentalfehler“ bezüglich  $Y$  und  $X$  sei als  $q$  bezeichnet. Nach DEMING<sup>11</sup> lassen sich aus folgender Gleichung fallspezifische Definitionen des Anstiegs ableiten.

$$\hat{a}_1^2 + \hat{a}_1 \left( \frac{q^2 SS_{XX} - SS_{YY}}{SS_{XY}} \right) - q^2 = 0 \quad (18)$$

mit  $q = \frac{\sigma_Y}{\sigma_X}$  und  $Q = \left( \frac{q^2 SS_{XX} - SS_{YY}}{SS_{XY}} \right)$

(Die Größe  $Q$  verwenden wir für die Fälle D und E, bei denen  $q$  einen konkreten numerischen Wert annimmt. Man kontrolliere bei den Ableitungen bzw. Berechnungen immer die Einheiten: Beispielsweise ist die Einheit von  $q$ :  $[q] = su / cu$ .)

11 vgl. C. HARTMANN, J. SMEYERS-VERBEKE, D. L. MASSART: „Problems in method-comparison studies“, *Analisis* 21(1993)125-132

Wir unterscheiden die folgenden Spezialfälle:

- (A) Annahme: Der Fehler in Y überwiegt.
- (B) Annahme: Der Fehler in X überwiegt.
- (C) Annahme: Die Empfindlichkeit entspricht dem Anstieg.
- (D) Vorwissen: Das Fehlerverhältnis beider Größen ist bekannt.
- (E) Vorwissen: Das Fehlerverhältnis ist gleich Eins.
- (F) Vorwissen: Der Fehler von Y oder von X ist bekannt.

### Fall (A)

Bei allen bisherigen Erörterungen wurde davon ausgegangen, dass der Messfehler der (Einstell-) Größe X im Vergleich zum Messfehler von Y vernachlässigbar ist, d. h.  $q \rightarrow \infty$  führte zur üblichen Schätzung der Kalibrierparameter (LS-, OLS-Schätzung). Der Schätzwert der Empfindlichkeit  $a_1$ ,  $\hat{a}_1(\text{OLS})$ , ergibt sich aus Gleichung (6).

### Fall (B)

Hier ist die Sachlage gerade umgekehrt, also „invers“, was wir durch den Index „INV“ andeuten.

Mit  $q \rightarrow 0$  liegt ein Fall für die *inverse Regression* vor, die im folgenden Abschnitt 2.4.2 besprochen wird.

Aus Gleichung (18) berechnet man schnell, dass dann

$$\hat{a}_1(\text{INV}) = \frac{SS_{YY}}{SS_{XY}} \quad (19)$$

### Fall (C)

Wenn der Anstieg dem Fehlerverhältnis entsprechen soll, wird das *Modell der reduzierten Hauptachse* (RHA) oder der Strukturrelation verwendet (engl.: reduced major axis, structural relationship, line of organic correlation). Mit  $q \rightarrow a_1$  ergibt sich aus Gleichung (18):

$$\hat{a}_1(\text{RHA}) = \text{sgn}(r_{XY}) \sqrt{\frac{SS_{YY}}{SS_{XX}}} \quad (20)$$

Dabei ist  $\text{sgn}(r_{XY})$  das Vorzeichen des Korrelationskoeffizienten der beiden Größen, der Maßzahl für den Zusammenhang der beiden Größen.

Man muss den Korrelationskoeffizienten nicht berechnen, wenn eindeutig ist, dass die Kalibriergerade einen positiven oder negativen Anstieg hat. Wenn das nicht eindeutig ist, ist die Analysenmethode vermutlich nicht kalibrierfähig.

### Fall (D)

Ist ein Erfahrungswert („a posteriori“-Wert) für das Fehlerverhältnis  $q$  bekannt, ergibt sich mit diesem

$$\hat{a}_1(APO) = -\frac{Q}{2} + \operatorname{sgn}(r_{XY}) \sqrt{\frac{Q^2}{4} + q^2} \quad (21)$$

### Fall (E)

Wenn der Experimentalfehler bei beiden Größen etwa gleich groß ist, kann das *Orthogonalmodell* (ORTH) zutreffen. Mit  $q = 1$  vereinfacht sich Gleichung (21) unwesentlich. Wir bezeichnen die Empfindlichkeit dann mit  $\hat{a}_1(\text{ORTH})$ .

### Fall (F)

Wenn es aus der Erfahrung zuverlässige Schätzwerte für die Fehler einer der Modellgrößen gibt, kann man sie wie folgt einbeziehen und damit  $\hat{a}_1(\text{OLS})$  modifizieren:

$$\hat{a}_1(APOX) = \frac{SS_{XY}}{SS_{XX} - N_C \sigma_X^2} \quad (22)$$

Gleichermaßen kann  $\hat{a}_1(\text{INV})$  realistischer berechnet werden durch:

$$\hat{a}_1(APOY) = \frac{SS_{YY} - N_C \sigma_Y^2}{SS_{XY}} \quad (23)$$

Mit ähnlichen Korrekturen soll die Schätzmethode von MANDEL arbeiten.<sup>12</sup>

❶ Datenbeispiel siehe Abschnitt 2.10.5

## 2.4.2 INVERSE KALIBRATION

Das inverse Kalibriermodell hat die Form:

$$X = a_0 + a_1 Y \quad (24)$$

12 vgl. C. HARTMANN, J. SMEYERS-VERBEKE, D. L. MASSART: „Problems in method-comparison studies“, *Analisis* 21(1993)125-132

Daraus ergibt sich der Schätzwert für die Empfindlichkeit

$$\hat{a}_1(IRG) = \frac{1}{\hat{a}_1(INV)} = \frac{SS_{XY}}{SS_{YY}} \quad (25)$$

Anwendungen der inversen Regression in der Analytik sind aus der Literatur schon lange bekannt.<sup>13</sup>

### 2.4.3 GEWICHTETE REGRESSION

Hierfür müssen Mehrfachmessungen jeder Kalibrierprobe bzw. von  $N_{REP}$  Proben vorliegen. Mit den entsprechenden Varianzen  $s^2(y_n)$  können dann die Messergebnisse gewichtet werden. Im einfachsten Fall kann der Wichtungsfaktor  $w$  nach

$$w_n^2 = \frac{1}{s^2(y_n)} \quad (26)$$

berechnet werden ( $n = 1, \dots, N_C$  oder  $n = 1, \dots, N_{REP}$ ).

Das „Gewicht“ an einer Messstelle ist also umso größer, je kleiner der dortige Messfehler ist. Beim Vorliegen von Wiederholungen der  $x_n$  an einer Kalibrierstelle  $x_n$  berücksichtigt die *iterativ gewichtete Regression* (manchmal auch engl.: maximum likelihood regression) explizit auch deren Varianz  $s^2(x_n)$  in geeigneter Weise.

Manche Autoren zählen diese Variante zu den robusten Techniken.

### 2.4.4 ROBUSTE REGRESSION

Die robuste Regression gehört zu den *parameterfreien Verfahren* und geht von den geordneten Einzelwerten aus:

$$y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(N)} \quad (27)$$

Somit können insgesamt  $N(N - 1)/2$  verschiedene Anstiege  $a_{1,ij}$  berechnet werden, wobei  $i < j$  gelten soll:

$$\hat{a}_{1,ij} = \frac{y_j - y_i}{x_j - x_i} \quad (28)$$

Man ordnet diese Koeffizienten dann der Größe nach und wählt aus der Reihe bei ungeradem  $N$  den Median und bei geradem  $N$  den Mittelwert aus den „beiden mittleren“ Werten als einen möglichen Schätzwert für den Anstieg aus:

13 vgl. S. D. OMAN: „Analyzing Residuals in Calibration Problems“, *Technometrics* 26(1984)347-353, A. DOBRIGAL, D. A. S. FRASER, R. GEBOTYS: „Linear calibration and conditional inference“, *Commun. Stat. - Theory & Meth.* 16(1987)1037-1048

$$\hat{a}_1 = \text{med}(\hat{a}_{1,ij}) \quad (29)$$

Von allen Absoluttermen

$$\hat{a}_{0,ij} = \frac{x_j y_i - x_i y_j}{x_j - x_i} \quad (30)$$

ist der Median ebenfalls ein möglicher Schätzwert für  $a_0$ .

Ausführlich beschrieben sind verschiedene Techniken der robusten Regression bei ROUSSEEUW und LEROY (1987). Bei DOERFFEL (1990, S. 174ff.) ist speziell für Methodenvergleiche die Variante von PASSING und BABLOK beispielhaft dargestellt. Auch DANZER (1989)<sup>14</sup> stellt detailliert ein Beispiel zur robusten Regression vor. Verschiedene Schätzverfahren vergleicht DIETZ<sup>15</sup> und gibt Empfehlungen, welche Variante bei bestimmten statistischen Situationen günstig ist.

Über Ergebnisse robuster Regression mit der IRLS-Technik (engl.: iteratively reweighted least squares) berichten bereits PHILLIPS und EYRING.<sup>16</sup> VANKEERBERGHEN et al.<sup>17</sup> wenden eine robuste Technik namens SRM (engl.: slope ranking method) an, um nichtlineares Signalverhalten in der Graphitrohr- AAS nachzuweisen.

❶ Datenbeispiel siehe Abschnitt 2.10.5

## 2.4.5 FUZZY REGRESSION

Wenn sowohl die Signal- als auch die Analytgröße fehlerbehaftet ist, kann anstelle der robusten Regression die Anwendung der *Theorie unscharfer Mengen* erwogen werden. Obwohl sie relativ aufwändig ist, stellen OTTO und BANDEMER<sup>18</sup> sie erstmals als alternative Kalibriermethode vor. Sie gehen von den hypothetischen Daten aus, die schon von PHILIPPS und EYRING (1983) verwendet wurden.

Der Algorithmus geht davon aus, dass jedem Datenpunkt, also einem Objekt  $(x_n, y_n)$  bzw.  $(x_n, \hat{y}_n)$  der Wert  $m_n(x_n, y_n)$  einer *membership function* zugeordnet werden kann, der zwischen Null und Eins liegen kann. Ist er gleich Null, so gehört der Kalibrierpunkt nicht zur Wertemenge der Kalibration, ist er gleich Eins, so gehört er zweifelsfrei dazu. In der Realität wird man infolge von Experimentalfehlern mit einer unsicheren oder unscharfen Zuordnung rechnen müssen. Diese Unschärfe, genauer die entsprechende Zugehörigkeitsfunktion wird nun modelliert.

14 K. DANZER: „Robuste Statistik in der analytischen Chemie“, Fresenius Z Anal. Chem. 335(1989)869-875

15 E. J. DIETZ: „A comparison of robust estimators in simple linear regression“, Commun. Stat. - Simul. 16(1987)1209-1227, E. J. DIETZ: „Teaching Regression in a Nonparametric Statistics Course“, The Amer. Statist. 43(1989)35-40

16 G. R. PHILLIPS, E. M. EYRING; „Comparison of Conventional and Robust Regression in Analysis of Chemical Data“, Anal. Chem. 55(1983)1134-1138

17 P. VANKEERBERGHEN, J. SMEYERS-VERBEKE, A. THIELEMANS, D. L. MASSART: „The slope ranking method, an algorithm for diagnosing linearity of calibration lines“, Analusis 20(1992)103-109

18 M. OTTO, H. BANDEMER: „Calibration with imprecise signals and concentrations based on fuzzy theory“, Chemom. Intell. Lab. Syst. 1(1986)71-78

Während OTTO und BANDEMER dafür einen kreisförmigen (paraboloiden) Ansatz wählen, gehen HU et al. (1990)<sup>19</sup> von einem linienförmigen Ansatz aus. Ihr Verfahren ähnelt damit einem gewichteten Modell.

Letztlich sollen derartige Methoden vor allem eine robuste Schätzung der Regressionsparameter und eine Ausreißererkenkung ermöglichen. Weniger aufwändige Methoden können daher zu gleichen Ergebnissen führen. Mit anderen Worten: Wenn der Datensatz ausreißerfrei ist, sollte man der üblichen linearen Regression den Vorzug geben. Deshalb weisen die letztgenannten Autoren anfangs ihres Beitrags auch zu Recht darauf hin, dass man zuallererst eine visuelle Inspektion der Daten vornehmen sollte, um Ausreißer zu erkennen.

Interessant für uns könnte ein speziell für die FAAS (Flammen-Atomabsorptionsspektroskopie, Graphitrohrtechnik) entwickelter Qualitätskoeffizient  $QC$  sein, dessen Werte unter 5% liegen sollten, wenn eine LS-Schätzung der Regressionsparameter „gut genug“ ist.

$$QC = \frac{100}{\sqrt{N_C - 1}} \sqrt{\sum_{n=1}^{N_C} \left( \frac{y_n - \hat{y}_n}{y_n} \right)^2} \quad (31)$$

Die gleiche Formel wurde als „Fehlerindex“ für Konzentrationswerte benutzt.<sup>20</sup>

## 2.5 Nichtlineare Kalibriermodelle

### 2.5.1 MODELLE 1. ART (QUASILINEARE)

In diesen Modellen erscheint die Einflussgröße  $X$  meist in Polynomgleichungen der Art

$$Y = a_0 + a_1 X + a_2 X^2 + a_3 X^3 + \dots \quad (32)$$

Solche Modelle heißen auch *quasilinear*, weil Potenzen von  $X$  durch einfache Substitution als neue Größen geschrieben werden können:

$$Y = a_0 + a_1 X_1 + a_2 X_2 + a_3 X_3 + \dots \quad (33)$$

19 Y. HU, J. SMEYERS-VERBEKE, D. L. MASSART: „An Algorithm for Fuzzy Linear Calibration“, Chemom. Intell. Lab. Syst. 8(1990)143-155

20 J. OPFERMANN, M.-L. MARTIN, M. HEYNISCH, K. DANZER: „Optimale Strategie zur Ermittlung von Eichkurven, dargestellt am Beispiel der AAS“, Fresenius Z. Anal. Chem. 328(1987)13-17

Das Modell ist damit wieder linear bezüglich der zu schätzenden Koeffizienten, und gängige Verfahren der OLS-Schätzung für den nun allerdings multivariaten Fall sind anwendbar.

Das bringt jedoch keine Probleme mit sich: Da eine Voraussetzung der Regression im multivariaten Fall ist, dass die Einflussgrößen nicht korreliert sein dürfen, können die einzelnen Koeffizienten unabhängig voneinander ermittelt werden, d. h. durch mehrmalige einfache lineare Regression.

❶ Datenbeispiel siehe Abschnitt 2.10.6

## 2.5.2 MODELLE 2. ART („ECHT NICHTLINEARE“)

In diesen Modellen ist wenigstens einer der Kalibrierkoeffizienten „nichtlinear“, wie zum Beispiel in

$$Y = a X^b \quad (34)$$

Auch in diesen Fällen können die Modelle (Gleichungen) durch geeignete Transformationen linearisiert werden. So wird aus der vorigen Gleichungen durch

$$Y' = \log_{10}(Y); \quad a_0 = \log_{10}(a); \quad a_1 = b; \quad X' = \log_{10}(X) \quad (35)$$

die linearisierte Gleichung

$$Y' = a_0 + a_1 X' \quad (36)$$

Allerdings können bei solchen Transformationen zum Beispiel die Varianzeigenschaften von  $Y$  und  $X$  dermaßen verändert werden, dass ein lineares Schätzverfahren für die Koeffizienten ausgeschlossen wird, vgl. dazu alternative Möglichkeiten (Abschnitt 2.2.2). So erzeugt zum Beispiel die Linearisierung von

$$Y = \frac{a_1 X}{a_0 + X} \quad (37)$$

mit

$$Y' = \frac{1}{Y}; \quad a_0' = \frac{1}{a_1}; \quad a_1' = \frac{a_0}{a_1}; \quad X' = \frac{1}{X} \quad (38)$$

einen heteroskedastischen Fall, d. h. eine Situation, wo die Varianz von  $Y$  über den Kalibrierbereich nicht konstant bleibt.

In der Literatur und in allgemein zugänglichen Rechenprogrammen existieren jedoch geeignete nichtlineare Schätzverfahren.

## 2.6 Beurteilung von Kalibriermodellen

### 2.6.1 BESTIMMTHEITSMAB ( $B$ )

Im Falle des einfachen linearen Regressionsmodells nach Gleichung (1) ist das *Bestimmtheitsmaß*  $B_{XY}$  (engl.: coefficient of determination) gleich dem Quadrat des Korrelationskoeffizienten  $r_{XY}$ .

Der Korrelationskoeffizient selbst sollte niemals zur Beurteilung der Güte eines empirischen Modells herangezogen werden.

$B_{XY}$  ist ein Maß für den durch das lineare Modell erfassten bzw. erklärten Anteil der Gesamtvarianz:

$$B_{XY} = \frac{\sum_{n=1}^{N_C} (\hat{y}_n - \bar{y})^2}{\sum_{n=1}^{N_C} (y_n - \bar{y})^2} \quad (39)$$

Mit der Testgröße

$$F_{EMP} (B_{XY}) = \frac{B_{XY} (N_C - 2)}{1 - B_{XY}} \quad (40)$$

wird gegen den kritischen Wert von  $F_{CRIT} (1, N_C - 2; 1 - \alpha)$  bzw.  $t_{CRIT}^2 (N_C - 2; 1 - \alpha/2)$  in der üblichen Weise getestet.

Damit ist zugleich der Test des linearen Korrelationskoeffizienten definiert.

Wenn der Test keine Signifikanz ergibt, sollte man ein nichtlineares Modell probieren.

❶ Datenbeispiel siehe Abschnitt 2.10.7

### 2.6.2 ANPASSUNGSGÜTE ( $GOF$ )

Die *Anpassungsgüte* ( $GOF$ , engl.: goodness of fit), ist ein Maß für den Anteil der Restvarianz, die durch das Modell erklärt wird. Als Testgröße dient:

$$F_{EMP} (GOF) = \frac{\sum_{n=1}^{N_C} (\hat{y}_n - \bar{y})^2}{\sum_{n=1}^{N_C} (y_n - \hat{y}_n)^2} \cdot \frac{N_C - N_P}{N_P - 1} \quad (41)$$

Ihr Wert wird gegen  $F_{CRIT} (N_P - 1, N_C - N_P; 1 - \alpha)$  getestet.

Auch dieser Test sollte Signifikanz anzeigen, anderenfalls sollte das lineare Modell verändert oder ein nichtlineares Modell verwendet werden. Der Test wird auch als Test der Signifikanz der Regression bezeichnet.

❶ Datenbeispiel siehe Abschnitt 2.10.7

### 2.6.3 ANPASSUNGSDEFEKT (LOF)

Erst mit Vorliegen von Mehrfachmessungen an jeder Kalibrierstelle oder wenigstens an  $N_{REP}$  Stellen des Kalibrierbereichs ist es möglich, den Anteil des experimentellen Fehlers an der Gesamtvarianz der Daten zu charakterisieren.

Am fundiertesten geschieht das im Rahmen der *Varianzanalyse*.

Zwei Testgrößen seien explizit und unabhängig vom Kontext der Varianzanalyse beschrieben. Der *Anpassungsdefekt* (LOF, engl.: lack of fit) beschreibt den Anteil des Modellfehlers am experimentellen Fehler.

Bei einer gleichen Anzahl von Wiederholungen ist seine Testgröße

$$F_{EMP} (LOF) = \frac{\sum_{n=1}^{N_{REP}} (\bar{y}_n - \hat{y}_n)^2}{\sum_{n=1}^{N_{REP}} \sum_{w=1}^{W_{REP}} (y_{nw} - \bar{y}_n)^2} \cdot \frac{N - N_{REP}}{N_{REP} - N_P} \quad (42)$$

$N$	Gesamtzahl aller Messungen (einschließlich der Mehrfachmessungen)
$N_{REP}$	Zahl der Kalibrierstellen, wo Mehrfachmessungen stattfanden
$W_{REP}$	Zahl der Mehrfachmessungen an diesen Kalibrierstellen
$N_P$	Zahl der Modellparameter (Koeffizienten)

Ihr Wert wird in der üblichen Weise mit  $F_{CRIT}(N_{REP} - N_P, N - N_{REP}; 1 - \alpha)$  verglichen.

Vor allem in der Versuchsplanung ist der *Adäquatheitstest* (TOA, engl.: test of adequacy) üblich, dessen Testgröße nach

$$F_{EMP} (TOA) = \frac{\sum_{n=1}^N (y_n - \hat{y}_n)^2}{\sum_{n=1}^{N_{REP}} \sum_{w=1}^{W_{REP}} (y_{nw} - \bar{y}_n)^2} \cdot \frac{N - N_{REP}}{N - N_P} \quad (43)$$

berechnet wird, und deren Wert mit  $F_{CRIT}(N - N_P, N - N_{REP}; 1 - \alpha)$  verglichen wird.

❶ Datenbeispiel siehe Abschnitt 2.10.8

## 2.7 Statistische Bewertung und Vertrauensbereiche

### 2.7.1 VARIANZANALYSE

Die Varianzanalyse ist eng verbunden mit der statistischen Versuchsplanung. Diese hat bisher für uns keine Rolle gespielt. Im Prinzip liegt aber unserem Kalibriermodell tatsächlich ein  $(N_c)^1$ -Versuchsplan zugrunde. Die Eins als Exponent codiert einen einfaktoriellen Versuchsplan, denn unser einziger Einflussfaktor ist die Konzentration  $X$ . Dieser Faktor ist in genau  $N_c$  Stufen, an unseren Kalibrierstellen, eingestellt worden.

Eine generelle statistische Hypothese lautet: Alle  $N_c$  Stichproben stammen aus derselben Grundgesamtheit und der  $X$ -Faktor hat keinen Einfluss auf  $Y$ .

Die Aufgabe der Varianzanalyse besteht nun darin, die Mittelwerte der  $N_c$  Versuchsgruppen miteinander zu vergleichen und somit die Wirkung des Faktors  $X$  auf  $Y$  zu bewerten.

In der Varianzanalyse werden ausschließlich die  $Y$ -Werte verwendet. Eine Bewertung der Modellierung der Abhängigkeit der  $Y$ -Werte von den  $X$ -Werten ist nicht deren Gegenstand. Lediglich die Information über die Stufenzahl  $N_c$  geht in die Analyse ein.

Der Grundgedanke der Varianzanalyse besteht nun in der Zerlegung der Gesamtstreuung der Messwerte von  $Y$  in einen faktor-, d. h. konzentrationsbedingten und einen zufälligen Anteil.

Genaugenommen wird mit Summen der jeweiligen Abweichungsquadrate (SAQ) gearbeitet bzw. den mittleren Quadratesummen (MAQ), den Quotienten aus den Quadratesummen SAQ und ihren Freiheitsgraden ( $fg$ ).

Solche Quadratsummen haben wir bereits verwendet, um Geradenanstiege zu berechnen. So können wir  $SS_{YY}$  zur Charakterisierung der „Gesamtstreuung“ verwenden.

Zur detaillierten Ausführung einer Varianzanalyse, ANOVA, empfehlen wir die Verwendung geeigneter Statistik-Software.

Die charakteristische Auswertetabelle einer einfaktoriellen Varianzanalyse sieht etwa so aus:

Varianzursache	SAQ	$fg =$	MAQ	$F_{EMP}$
zwischen $X$ -Stufen	$SS_{ZWI}$	$fg_{ZWI} = N_c - 1$	$MAQ_{ZWI} = SS_{ZWI} / fg_{ZWI}$	$MAQ_{ZWI} / MAQ_{INN}$
innerhalb $X$ -Stufen	$SS_{INN}$	$fg_{INN} = N - N_c$	$MAQ_{INN} = SS_{INN} / fg_{INN}$	
total	$SS_{TOT} = SS_{YY}$	$fg_{TOT} = N - 1$		

❶ Datenbeispiel siehe Abschnitt 2.10.9

Kommen wir nun zu varianzanalytischen Aspekten, nachdem wir eine Signalmodellierung vorgenommen haben.

Liegen an den Stellen  $x_n$  des Kalibrierbereichs lediglich *Einzelmessungen*  $y_n$  vor, kann die Güte des Kalibriermodells nur pauschal als Summe von systematischem Modell- und

zufälligem Experimentalfehler bewertet werden. (Man vergleiche die Tests des Bestimmtheitsmaßes  $B_{xy}$  und der Anpassungsgüte  $GOF$ .)

Im varianzanalytischen Sinn kann die Quadratesumme der zentrierten Messwerte,  $SS_{yy}$ , die der Gesamtvarianz entspricht, vorerst nur in zwei Anteile zerlegt werden:

$$\sum_{n=1}^{N_C} (y_n - \bar{y})^2 = \sum_{n=1}^{N_C} (\hat{y}_n - \bar{y})^2 + \sum_{n=1}^{N_C} (y_n - \hat{y}_n)^2 \quad (44)$$

Der erste Summand (nennen wir ihn:  $SS_{MOD}$ ) geht auf die durch das Kalibriermodell erklärte Varianz und der zweite Summand (nennen wir ihn:  $SS_{RES}$ ) geht auf den experimentellen Zufallsfehler zurück.

Mit dem ersten Summanden sind  $N_p - 1$ , mit dem zweiten Summanden  $N_C - N_p$  Freiheitsgrade verknüpft, die zusammen  $N_C - 1$  ergeben, die Freiheitsgrade des Kalibrierexperiments.

Man stellt die mittleren Abweichungsquadrate der beiden Summanden als „Varianzanteile“ in der schon als  $GOF$ -Test verwendeten Testgröße

$$F_{EMP} = \frac{\sum_{n=1}^{N_C} (\hat{y}_n - \bar{y})^2}{\sum_{n=1}^{N_C} (y_n - \hat{y}_n)^2} \cdot \frac{(N_C - N_p)}{(N_p - 1)} \quad (45)$$

gegenüber und testet deren Wert gegen  $F_{CRIT}(1, N_C - 2; 1 - \alpha)$ , sofern eine Kalibriermodell nach Gleichung (1) verwendet wurde.

Die statistische Nullhypothese lautet hier in anderer Formulierung „Das Modell ist nicht vom Typ Gleichung (1)“. Man entscheidet sich bei signifikanter Abweichung, dem Zutreffen der Alternativhypothese, dafür, ein Modell wie in Gleichung (1) anzunehmen. Im Prinzip gilt der Test für ein beliebiges Modell mit zwei Parametern.

Zur Durchführung einer regressionsbezogenen Varianzanalyse müssen jedoch für mindestens eine Stelle  $x_n$  *Wiederholungsmessungen* von  $Y$  vorliegen. Dann erst kann man Modell- und Experimentalfehler detailliert und statistisch sicher gegeneinander abwägen.

Auch das ist uns nicht neu: Im Abschnitt 2.6.3 hatten wir entsprechende Vergleichstests bereits kennen gelernt, den  $LOF$ -Test und den Adäquatheitstest ( $TOA$ ).

❶ Datenbeispiel siehe Abschnitt 2.10.9

### 2.7.2 VERTRAUENSBEREICH DER SIGNALWERTE IN DER KALIBRIERPHASE

Im Kalibrierschritt kann zu jedem einzelnen  $\hat{y}_n$ , also zu allen zu  $x_n$  gehörenden Punkten, mit der in Gleichung (11) definierten Standardabweichung ein Vertrauensbereich angegeben werden:

$$cnf(\hat{y}_n) = \hat{y}_n \pm t(fg; q) s(\hat{y}_n) \quad (46)$$

$$cnf(\hat{y}_n) = \hat{y}_n \pm t(fg; q) s_{RES} \sqrt{\frac{1}{N_C} + \frac{(x_n - \bar{x}_C)^2}{SS_{XX}}} \quad (47)$$

Darin sind  $fg = N_C - 2$  und  $q = 1 - \alpha/2$ .<sup>21</sup>

Der Kalibriergraph nimmt damit seine charakteristische Form an.

Innerhalb des Konfidenzbandes liegen nun alle Vorhersagewerte  $\hat{y}_n$  mit einer Wahrscheinlichkeit von  $q$  bzw. einem Irrtumsrisiko von  $\alpha$ .

Man kann bereits aus Gleichung (47) ableiten, worin der Vorteil der Erhöhung von  $N_C$  bestünde. Infolge von mehr Kalibrierstellen oder von Wiederholungsmessungen im Kalibrierschritt kann die Breite des Vertrauensbandes um den Kalibriergraphen verringert werden.

❶ Abbildung (Kalibriergraph mit Konfidenzband) im Datenbeispiel Abschnitt 2.10.8

### 2.7.3 VERTRAUENSBEREICH DER SIGNALWERTE IN DER ANWENDUNGSPHASE

Unter Berücksichtigung einer Einzelmessung im Anwendungsschritt wird der Fehler der  $\hat{y}_n$  etwas größer:

$$s(\hat{y}_n) = s_{RES} \sqrt{1 + \frac{1}{N_C} + \frac{(x_n - \bar{x}_C)^2}{SS_{XX}}} \quad (48)$$

und das Vertrauensintervall vergrößert sich:

$$cnf(\hat{y}_n) = \hat{y}_n \pm t(fg; q) s_{RES} \sqrt{1 + \frac{1}{N_C} + \frac{(x_n - \bar{x}_C)^2}{SS_{XX}}} \quad (49)$$

<sup>21</sup> Aus der Formel geht prinzipiell das am Anfang und Ende des Kalibrierbereichs aufgeweitete Konfidenzband hervor. Andererseits gilt nach C. HARTMANN, J. SMEYERS-VERBEKE, D. L. MASSART, *Analisis* 21(1993)125-132: „It is known that the most precise estimates of the regression coefficients are obtained when most data points are located at the extremes of the range considered.“

Sieht man für den Anwendungsschritt  $N_A$  Messungen des Analytsignals vor, so verringert sich dieser Fehler zufolge

$$s(\hat{y}_n) = s_{RES} \sqrt{\frac{1}{N_A} + \frac{1}{N_C} + \frac{(x_n - \bar{x}_C)^2}{SS_{XX}}} \quad (50)$$

und das Vertrauensintervall ändert sich entsprechend. Freiheitsgrad und Ordnung des Quantils der  $t$ -Verteilung bleiben unverändert.

Quadriert man die Gleichungen (48) bzw. (50) und berücksichtigt Gleichung (13), so ergibt sich jeweils die Varianz der Signalwerte  $\hat{y}_n$  als Summe aus zwei Anteilen, die einerseits von  $s_{RES}^2$  und andererseits von  $s^2(\hat{a}_1)$  abhängen.

#### 2.7.4 VERTRAUENSBEREICH EINES ANALYSENERGEBNISSES

Nach EBEL<sup>22</sup> hat man als Vertrauensintervall für einen mittleren Analytwert in der Analysenprobe mit

$$cnf(\bar{x}_A) = \bar{x}_A \pm \frac{s_{RES}}{a_1} \sqrt{\frac{t^2(fg_A; q)}{N_A} + \frac{t^2(fg_C; q)}{N_C} + \frac{t^2(fg_C; q)(\bar{x}_A - \bar{x}_C)^2}{SS_{XX}}} \quad (51)$$

zu rechnen, was sich für „ähnliche“ Freiheitsgrade in den  $t$ -Werten zu

$$cnf(\bar{x}_A) = \bar{x}_A \pm \frac{s_{RES} t(fg; q)}{a_1} \sqrt{\frac{1}{N_A} + \frac{1}{N_C} + \frac{(\bar{x}_A - \bar{x}_C)^2}{SS_{XX}}} \quad (52)$$

vereinfacht.

Mit der Beziehung

$$\bar{y}_A - \bar{y}_C = a_1 (\bar{x}_A - \bar{x}_C) \quad (53)$$

lässt sich Gleichung (52) auch anders darstellen.

Aus Gleichung (52) lassen sich Möglichkeiten zur Verkleinerung des Vertrauensintervalls ableiten, etwa sollte  $N_A > 1$  sein, möglichst sei  $N_A = N_C$ , am besten  $N_C \geq 5$ .

22 S. EBEL: „Über den Vertrauensbereich kalibrierender Analysenverfahren; Teil 1: Lineare Kalibrierfunktionen“, CAL 1(1983)55-6; Siehe dazu evtl. auch: S. EBEL „Fehler und Vertrauensbereiche analytischer Ergebnisse“, in: Analytiker-Taschenbuch, Band 11, Springer, Berlin, 1993, 3-59 oder S. EBEL: „Kalibrierung und Auswertung in der Analytik“, in: H. NAUMER, W. HELLER (Hrsg.): „Untersuchungsmethoden in der Chemie“, 3. erweit. u. neubearb. Aufl., Thieme, Stuttgart, 1997, Kap. 30, S. 471-474.

Nach MILLER<sup>23</sup> darf Gleichung (52) unter bestimmten Bedingungen auch für eine Einzelmessung der Testprobe, d. h.  $N_A = 1$ , benutzt werden.

❶ Datenbeispiel siehe Abschnitt 2.10.10

## 2.8 Qualitätssicherung und Validieren von Kalibriermodellen

### 2.8.1 STANDARDREFERENZMATERIALIEN

Das vielleicht größte Problem, vor allem bei Spurenanalysen, ist die Sicherung der Richtigkeit der Ergebnisse, d. h. das Vermeiden systematischer Fehler.

Systematische Fehlerquellen sind in jedem Schritt eines analytischen Prozesses möglich.

Die sicherste Methode zum Erkennen systematischer Fehler ist eine kontinuierliche und vielseitige *interne Qualitätssicherung*, insbesondere durch gelegentliche Analyse von (zertifizierten) *Standardreferenzmaterialien* (CRM, SRM) mit bekanntem Analytgehalt  $X_{REF}$ . Strenggenommen ist ein Analysenverfahren dann nicht kalibrierfähig, wenn geeignete, d. h. repräsentative, die Matrix der zu erwartenden Analysenproben adäquat repräsentierende, Standardreferenzmaterialien fehlen.

Man wird dann jedoch meist laborinterne Referenzproben herstellen können, deren Matrix weitgehend der Matrix der zu erwartenden Messproben ähnelt.

Andererseits bietet sich bei Problemen mit der Präparation solcher Referenzproben die Standard-Additionsmethode (SAM) an, bei der laborinterne Analytstandards schrittweise der Analysenprobe (Analyt und Matrix) zugegeben werden und man ständig die Summe beider Analyte (aus Probe und Standard) erfasst. (zur SAM siehe Abschnitt 4.3.1)

$$H_0: \bar{x}_{A,LABOR} = \bar{x}_{A,REF} \quad (54)$$

Entsprechend der Hypothese  $H_0$  lassen sich die gefundenen (mittleren) Analytwerte  $\bar{x}_{A,LABOR}$  mit den erwarteten  $\bar{x}_{A,REF}$  vergleichen. Die erforderlichen Tests und ihre Voraussetzungen sind bekannt.

Interne Qualitätssicherung kann auch durch *Methodenvergleiche* erfolgen.

Hierbei sollte man auf das gleiche analytische Problem möglichst zwei vom analytischen Prinzip völlig unterschiedliche (unabhängige) Methoden bzw. Verfahren anwenden. Am günstigsten ist es, die eigene Methode A ( $X_{LABOR}$ ) mit einer zertifizierten Methode B ( $X_{REF}$ ) zu vergleichen. Wenigstens akkreditierte Laboratorien halten solche Methoden vor.

---

23 J. N. MILLER: „Calibration Methods in Spectroscopy. V. Further Errors in Calibration Graphs“, Spectrosc. Int. 4(1992) no. 1, 41-43.

Trägt man die Ergebnisse beider Methoden in einem Streudiagramm gegeneinander auf, so liegen alle Werte im Idealfall auf einer Geraden mit der Steigung  $a_1 = 1$  und einem Achsenabschnitt  $a_0 = 0$  (YOUDEN-Plot<sup>24</sup>).

Die Modellgleichung ist:

$$X_A = a_0 + a_1 X_B \quad (55)$$

Man berechnet in Abhängigkeit vom Verhältnis des Zufallsfehlers der verglichenen Methoden, wie oben dargestellt, die den aktuellen Daten tatsächlich entsprechenden Geradenkenngrößen.

Sie werden dann als *systematischer additiver* ( $a_0$ ) bzw. *multiplikativer* oder *proportionaler* ( $a_1$ ) „Fehler“ interpretiert und getestet („statistisch gesichert“).

Ein solcher Vergleich kann prinzipiell auch Teil eines Verfahrens zur Prüfung der Gleichwertigkeit von Analysenverfahren sein.<sup>25</sup>

Wenn nicht lediglich die Vorhersagewerte der Analytgrößen gefragt sind, ist es auch möglich, die zugrundeliegenden Kalibriermodelle (für  $X_{\text{LABOR}}$  und  $X_{\text{REF}}$  bzw. für  $X_A$  und  $X_B$ ) selbst zu vergleichen. (Siehe dazu Abschnitt 2.9.1.)

Auch eine *externe Qualitätssicherung*, meist durch Teilnahme an Ringversuchen, trägt zum Aufdecken und Vermeiden systematischer Fehler bei.

Allerdings bietet die bloße Verwendung von SRM noch keine Garantie für eine richtige, fehlerfreie Kalibrierung. So sind zum Beispiel manche feste SRM für die Kalibrierung von Festkörper-Mikroanalysen wertlos, wenn man nicht die Gefahr von Mikroinhomogenitäten erkennt und die minimale Messprobenmenge entsprechend wählt.<sup>26</sup>

Dies illustriert nicht nur die Schwierigkeit, „perfekte“ Qualitätssicherung zu betreiben, sondern auch noch einmal die Probleme zuverlässiger Spurenanalytik überhaupt.

Man beachte die aktuelle internationale Normung zum Thema „Lineare Kalibrierung mit Referenzmaterialien“.<sup>27</sup>

## 2.8.2 VALIDIERUNG

Obwohl vor allem bei den multivariaten, d. h. den faktoriellen Methoden, die Aufgabe der Validierung besteht, könnte prinzipiell auch bei univariaten Methoden die Frage nach einem Gütemaß gestellt werden, dass die Prognosefähigkeit des Modells charakterisiert.

24 W. J. YOUDEN: „Statistical Methods for Chemists“, Wiley, New York, 1955, S. 41.

25 vgl. DIN 38 402, Teil 71 (1987): „DEV zur Wasser-, Abwasser- und Schlammuntersuchung (Gruppe A) Gleichwertigkeit zweier Analysenverfahren aufgrund des Vergleiches der Untersuchungsergebnisse an der gleichen Probe“, Beuth Verlag, Berlin, Köln, zitiert nach: G. BRAUN: „Gleichwertigkeit von Analysenverfahren“, GIT Labor-Fachzeitschrift, Heft 2 (2000)116-118.

26 J. PAUWELS, C. VANDECASTEELE: „Determination of the minimum sample mass of a solid CRM to be used in chemical analysis“, Fresenius J. Anal. Chem. 345(1993)121-123: „most CRMs should not be used to calibrate as it is impossible to evaluate the uncertainty to be assigned to the certified value if samples smaller than the ‘minimum sample size’ are used“.

27 Im Juli 1993 fand man sie unter DIN ISO 11 095, Beuth Verlag, Berlin, Köln.

Zum Testen des aufgestellten Kalibriermodells benötigt man dann unabhängige Kalibrierproben, die *Validierproben*, die nicht zur Modellierung verwendet wurden. Für die Validierproben müssen die Werte der Analytgröße ebenfalls bekannt sein und dienen nun als Referenzwerte.

Hat man keine echten, d. h. zusätzlichen Validierproben, so kann die Technik der *Kreuzvalidierung* angewendet werden. Dabei wird aus der Gesamtzahl der Kalibrierproben eine kleine Teilmenge als Validierprobenmenge reserviert und mit den restlichen Proben das Kalibriermodell erstellt. Danach wird der Modellfehler mit den reservierten Proben berechnet.

Die Validierproben werden nun wieder zur Menge der Kalibrierproben gegeben, und es wird eine neue Teilmenge zum Zwecke des Validierens ausgewählt usw.

Dieser Vorgang kann je nach Strategie des Validierens nun einige Male wiederholt werden.

Es gibt zahlreiche Varianten des *Re-Samplings* zur Erzeugung neuer Lern-Stichproben, die auch zur besseren Schätzung der Kalibrierparameter eingesetzt werden.

In der Diskriminanzanalyse (vgl. Abschnitt 6.3.2) wurde von LACHENBRUCH die (verteilungsfreie) U-Methode eingeführt, die auch als zyklisches Austauschverfahren, „hold out“, „leaving one out“, „cross validation“ oder „jackknifing“ bezeichnet wird. Bei ihr wird in zyklischer Weise jeweils ein Objekt der Lernmenge (Kalibrierprobe) bei der Berechnung der Diskriminanzfunktionen weggelassen und dann die Fehlereinschätzung mit diesem Objekt vorgenommen. Diese Methode wurde dahingehend modifiziert, dass gleichzeitig mehrere Objekte ausgelassen werden. Dadurch verringert sich zwar der Aufwand, aber auch die Zuverlässigkeit der Schätzung. Eine andere Methode des jackknifing, die *Bootstrap-Methode* nach HARTIGAN wurde so modifiziert, dass eine Kreuzvalidierung des gesamten Verfahrens ermöglicht wird. Dabei wird eine Zufallsstichprobe gleichen Umfangs wie die Lernmenge aus dieser gezogen und zum Modelltest die volle Lernmenge als Testmenge benutzt. Bei kleinen Stichproben liefert allerdings auch diese Methode zu optimistische Schätzungen.

Entsprechend Gleichung (54) können die Referenzwerte aus einer chemischen Referenzmethode den z. B. mittels NIR (Nahinfrarot-Spektroskopie) gefundenen Analytwerten gegenübergestellt werden. Der Zusammenhang beider Wertevektoren darf dann durch den Korrelationskoeffizienten  $r$  charakterisiert werden.

Dieser ist wie folgt definiert und entspricht damit der Definition des Bestimmtheitsmaßes für die Signaldomäne in Gleichung (39).

$$r = \sqrt{\frac{\sum_{n=1}^N (\hat{x}_{NIR,n} - \bar{x}_{REF})^2}{\sum_{n=1}^N (x_{REF,n} - \bar{x}_{REF})^2}} \quad (56)$$

Die Summe der quadrierten Differenzen aus Vorhersage- und Referenzwerten wird **PRESS** (engl.: predictive sum of squares) genannt:

$$PRESS = \sum_{n=1}^N (\hat{x}_{NIR, n} - x_{REF, n})^2 \quad (57)$$

Man erkennt die Analogie zum Zähler der (Rest-)Varianz der Regression, den wir oben kennen gelernt haben.

*PRESS* lässt sich generell für Abweichungen errechneter von vorgegebenen Werten definieren.

Daraus lassen sich zwei Typen von „Restfehlern“ ableiten.

Aus *PRESS* erhält man mit

$$N = N_{CAL}$$

$$N = N_{VAL}$$

d. h. Validierproben sind  
(einige der) Kalibrierproben

d. h. Validierproben sind  
unabhängig von Kalibrierproben

den

*Kalibrierfehler, Modellfehler, SEC.*

*Validierfehler, Vorhersagefehler, SEP.*

Je nach berücksichtigtem Freiheitsgrad *fg* existiert eine Vielzahl weiterer Gütemaße bzw. Begriffe, die sich aus *PRESS* ableiten:

$$TE, SEC, RMSD, SEE, SEP = \sqrt{\frac{PRESS}{fg}} \quad (58)$$

Bei <i>TE</i>	ist	$fg = 1,$
bei <i>SEC</i>	ist	$fg = N_{CAL} - N_{PAR}$ oder $N_{CAL} - N_{FACT}$ oder $N_{CAL} - 1,$
bei <i>RMSD</i>	ist	$fg = N_{CAL},$
bei <i>SEE</i>	ist	$fg = N_{CAL} - 1,$
bei <i>SEP</i>	ist	$fg = N_{VAL} - 1.$
Mit <i>SECV</i>	ist	der Fehler nach der oben geschilderten Methode der Kreuzvalidierung gemeint.

$N_{PAR}$  ist die Zahl der Parameter in der Regressionsfunktion, auf der letztlich die Werte  $\hat{x}$  beruhen.

$N_{FACT}$  ist die Zahl der Kalibrierfaktoren bei den multivariaten faktoriellen Methoden.

## 2.9 Vergleich von Kalibrier- bzw. Regressionsmodellen

### 2.9.1 VERGLEICH DER REGRESSIONSPARAMETER

Manchmal ist es erforderlich, zwei Kalibrier- bzw. Regressionsmodelle miteinander zu vergleichen. So muss man gelegentlich alternative Bestimmungsmethoden laborintern vergleichen oder sich mit Analysemethoden anderer Labors auseinandersetzen.

Im Vordergrund steht dabei der Test der Empfindlichkeiten.

Das statistische Vorgehen beim Vergleich von Regressionskoeffizienten ähnelt einem Mittelwertvergleich, dem immer ein Vergleich von Streumaßen vorausgeht. Wir empfehlen, im Bedarfsfalle geeignete Software zu verwenden oder die folgenden drei Vorgehensschritte selbst zu programmieren. Details dazu findet man zum Beispiel bei STORM (1995, S. 253-257).

Prinzipiell ist folgendermaßen vorzugehen:

- (1) Zuerst sind die Reststreuungen  $s_{\text{RES}}^2$  der alternativen Modelle mittels  $F$ -Test zu vergleichen. Vom Ausgang des  $F$ -Tests hängt die konkrete Form des anschließenden („doppelten“)  $t$ -Tests ab.
- (2) Wurde die Hypothese der Gleichheit beider Streuungen bestätigt, kann der  $t$ -Test in der üblichen Form erfolgen. Wird diese Hypothese allerdings verworfen, ist man auf ein Näherungsverfahren zum Vergleich der beiden Anstiege angewiesen (vgl. STORM).
- (3) Abschließend werden die beiden Regressionskonstanten (Absolutterme) ebenfalls mittels eines  $t$ -Tests verglichen.

❶ Datenbeispiel siehe Abschnitt 2.10.11

### 2.9.2 VERGLEICH DER GÜTEMAßE UND LEISTUNGSPARAMETER

Ein Vergleich ist über Gütemaße wie das Bestimmtheitsmaß möglich.

Es können aber auch die im vorigen Abschnitt besprochenen Maßzahlen  $SEC$ ,  $SEP$  oder ähnliche verwendet werden.

### 2.9.3 VERGLEICH VON REGRESSIONSMETHODEN

Eine Übersicht über verschiedene Regressionsmethoden findet man zum Beispiel bei HWANG und WINEFORDNER.<sup>28</sup>

---

28 J. D. HWANG, J. D. WINEFORDNER: „Regression methods in analytical chemistry“, Prog. Anal. Spectrosc. 11(1988)209-249.

HU et al.<sup>29</sup> vergleichen insgesamt vier Methoden der robusten und fuzzy Kalibrierung mit der LS-Methode (MKQ) insbesondere unter dem Aspekt der Ausreißerererkennung. Mit dem gleichen Ziel vergleicht auch WALCZAK<sup>30</sup> eine Reihe von robusten Methoden mit einer neuen Methode der Teilmengenauswahl.

Auf kleine Datensätze (4 bis 10 Wertepaare) haben RUTAN und CARR<sup>31</sup> fünf verschiedene Algorithmen, einschließlich zweier Methoden mit KALMAN-Filtern angewandt und u. a. bezüglich Ausreißerresistenz verglichen.

Bei Analysenmethoden-Vergleichen muss beachtet werden, dass beide Regressionsgrößen fehlerbehaftet sind. Anhand zweier Datensätze (normal- und rechteckverteilt) vergleichen HARTMANN et al.<sup>32</sup> Regressionsergebnisse nach OLS und der Methode von MANDEL.

Auf das Problem gleichzeitig fehlerhafter X- und Y-Variablen sowie der Korrelation von Absolutterm und Anstieg weisen auch GONZÁLEZ und ASUERO<sup>33</sup> hin und bieten ein Programm an zur Berechnung der Koeffizienten nach der MLH- und der EJCR-Methode (MLH = maximum likelihood, EJCR = elliptic joint confidence region). PHILLIPS und EYRING<sup>34</sup> vergleichen die konventionelle mit der iterativ gewichteten (robusten) Regression (IRLS).

---

29 Y. HU, J. SMEYERS-VERBEKE, D. L. MASSART: „Outlier Detection in Calibration“, Chemom. Intell. Lab. Syst. 9(1990)31-44.

30 B. WALCZAK: „Outlier detection in multivariate calibration“, Chemom. Intell. Lab. Syst. 28(1995)259-272.

31 S. C. RUTAN, P. W. CARR: „Comparison of robust regression methods based on least-median and adaptive KALMAN filtering approaches applied to linear calibration data“, Anal. Chim. Acta 215(1988)131-142.

32 C. HARTMANN, J. SMEYERS-VERBEKE, D. L. MASSART: „Problems in method-comparison studies“, Analisis 21(1993)125-132.

33 A.G. GONZÁLEZ, A. G. ASUERO: „Computational program for validating analytical methods“, Fresenius J. Anal. Chem. 346(1993)885-887.

34 G. R. PHILLIPS, E. M. EYRING: „Comparison of Conventional and Robust Regression in Analysis of Chemical Data“, Anal. Chem. 55(1983)1134-1138.

## 2.10 Datenbeispiele<sup>35</sup>

### 2.10.1 DATENBEISPIEL ZU 2.2.1 - KALIBRIERDATEN (EINFACHMESSUNGEN)

Wir machen in diesem Beispiel mit Grundbegriffen und Rechengrößen vertraut, die in Regressionsrechnungen verwendet werden. Dazu verwenden wir, vor allem zu Vergleichszwecken, einen Datensatz von MILLER (1991). Vor dem Beginn jeglicher Berechnungen sollte allerdings immer eine einfache visuelle Beurteilung der Datensätze stehen. Es ist geradezu ein handwerklicher Fehler, wenn man dies unterlässt. Wir gehen darauf in Abschnitt 2.10.3 ausführlicher ein.

❶ Daten-ID: J. N. MILLER: „Calibration Methods in Spectroscopy. I.“ Spectrosc. Int. 3(1991) no. 4, 41-43

Wir gehen zunächst vom einfachsten Fall aus, in dem nur Einfachmessungen an den Kalibrierpunkten vorliegen, die durch folgende Werte  $x$  für die Konzentration  $X$  und  $y$  für das spektroskopische Signal  $Y$  charakterisiert sind.

Die Einheiten der Größen seien allgemein bezeichnet mit  $[X] = cu$  und  $[Y] = su$ , wobei  $cu = „concentration unit“$  und  $su = „signal unit“$  bedeutet. In späteren Beispielen werden wir sie nicht explizit angeben.

Spalte 1	Spalte 2	Spalte 3
$n$	$x$	$y$
1	0	0,1
2	1	3,8
3	2	10,0
4	3	14,4
5	4	20,7
6	5	26,9
7	6	29,1

Eine grafische Darstellung zeigt keine großen Auffälligkeiten im Datenbeispiel. Auch liegen alle Datenpunkte innerhalb des Konfidenzbandes.

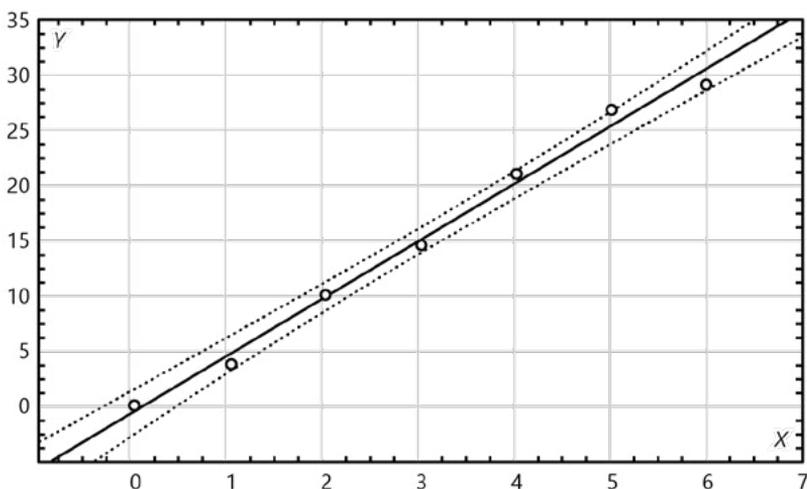


Abb. 1: Linearer Kalibriergraph mit Konfidenzband

35 In der Regel werden alle unabhängigen, gegebenen, übernommenen ( $x$ -)Werte *kursiv* gesetzt.

Für dieses Datenbeispiel erhalten wir im Einzelnen die folgenden Rechenwerte:

1	2	3	4	5	6	7	8	9
<b>n</b>	<b>x</b>	<b>y</b>	<b>x*y</b>	<b>x*x</b>	<b>y*y</b>	<b>x-<math>\bar{x}</math></b>	<b>y-<math>\bar{y}</math></b>	<b>(x-<math>\bar{x}</math>)(y-<math>\bar{y}</math>)</b>
1	0	0,1	0,0	0	0,01	-3	-14,9	44,7
2	1	3,8	3,8	1	14,40	-2	-11,2	22,4
3	2	10,0	20,0	4	100,00	-1	-5,0	5,0
4	3	14,4	43,2	9	207,40	0	-0,6	0,0
5	4	20,7	82,8	16	428,50	1	5,7	5,7
6	5	26,9	134,5	25	723,60	2	11,9	23,8
7	6	29,1	174,6	36	846,80	3	14,1	42,3
<b>N=</b>	<b>S<sub>x</sub> =</b>	<b>S<sub>y</sub> =</b>	<b>S<sub>xy</sub> =</b>	<b>S<sub>xx</sub> =</b>	<b>S<sub>yy</sub> =</b>	<b>SS<sub>xx</sub> =</b>	<b>SS<sub>yy</sub> =</b>	<b>SS<sub>xy</sub> =</b>
7	21	105	458,9	91	2320,7	28	745,7	143,9

Damit erhalten wir  $\hat{a}_1 = 5,139 \text{ su/cu}$  und, mit den Mittelwerten  $\bar{x} = 21/7 = 3$  und  $\bar{y} = 105/7 = 15$ , schließlich auch  $\hat{a}_0 = - 0,418 \text{ su}$ .

Nun können wir die Residuen berechnen, die wir für die Ableitung der signalbezogenen Kenngrößen benötigen.

1	2	3	10	11	12
<b>n</b>	<b>x</b>	<b>y</b>	<b><math>\hat{y}</math></b>	<b><math>y_{\text{RES}}</math></b>	<b><math>y_{\text{RES}}^2</math></b>
1	0	0,1	- 0,418	0,518	0,268
2	1	3,8	4,721	- 0,921	0,848
3	2	10,0	9,861	0,139	0,019
4	3	14,4	15,000	- 0,600	0,360
5	4	20,7	20,139	0,561	0,315
6	5	26,9	25,279	1,621	2,628
7	6	29,1	30,418	- 1,318	1,737
<b>N = 7</b>				<b><math>\Sigma = 0,000</math></b>	<b><math>SS_{\text{RES}} = 6,175</math></b>

Man beachte, dass die Summe  $\Sigma$  der Residuen Null beträgt.

## 2.10.2 DATENBEISPIEL ZU 2.2.2 - STATISTISCHE VORAUSSETZUNGEN

- (1) Die Größe  $X$  lässt sich an den Stellen  $x_n$  fehlerfrei einstellen. Das ist für unser Beispiel offensichtlich gegeben.
- (2) Die Größe  $Y$  ist an den Stellen  $x_n$  normalverteilt.

In Abschnitt 2.10.8 haben wir Wiederholungsmessungen simuliert. Mit diesen Daten können wir nun eine visuelle Bewertung dieser Voraussetzung vornehmen.

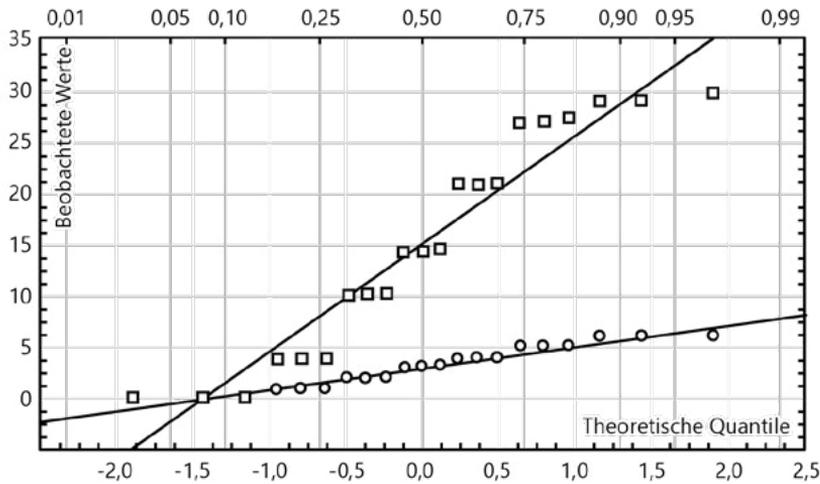


Abb. 2: Q-Q-Plot (unten für  $X$ , oben für  $Y$ )

In der Literatur finden sich hierzu auch kritische Werte von Testgrößen.

- (3) Die Varianz von  $Y$  an den Stellen  $x_n$  ist im Kalibrierbereich konstant.  
 Dies kann erst nach Vorliegen von Mehrfachmessungen geprüft werden (siehe Abschnitt 2.10.8). Zur Illustration haben wir für die folgende Abbildung die Messwerte der oberen Kalibrierstellen künstlich aufgeweitet.

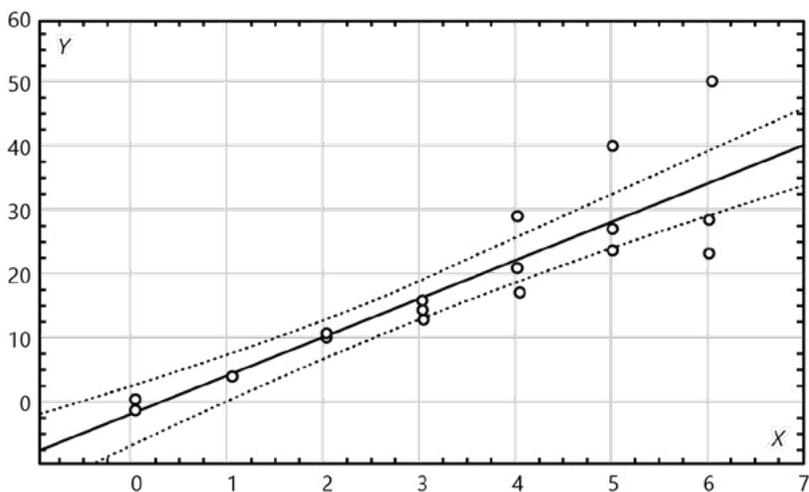


Abb. 3: Typische heteroskedastische Datenlage

- (4) Die Größe  $Y$  ist an den Stellen  $x_n$  ausreißerfrei.  
 Dies kann sinnvoll erst nach Vorliegen von Mehrfachmessungen statistisch geprüft werden (siehe Abschnitt 2.10.3).  
 Wie unterschiedlich sich Ausreißerfehler im Regressor-Bereich und im Regressanden-Bereich auswirken können, illustrieren wir anhand folgender Abbildungen.

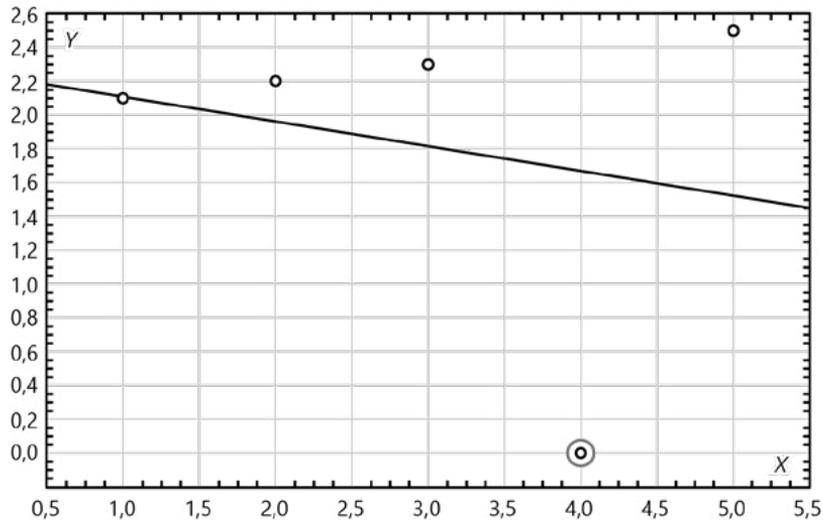


Abb. 4: Auswirkung eines Ausreißers im Y-Bereich

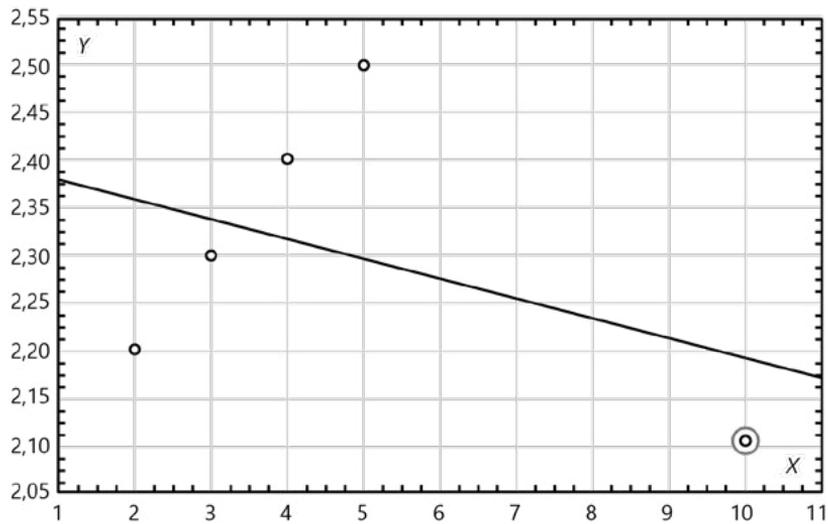


Abb. 5: Auswirkung eines Ausreißers in X-Bereich

- (5) Die Residuen  $y_{\text{RES}, n}$  sind normalverteilt und korrelieren nicht mit den  $x_n$ . Dies scheint in unserem Beispiel erfüllt zu sein. Alle heute üblichen Programme für Regressionsrechnungen bieten hierfür entsprechende grafische Darstellungen an. Mit den Mehrfachmessungen (Daten siehe Abschnitt 2.10.8) erhalten wir das folgende Bild.

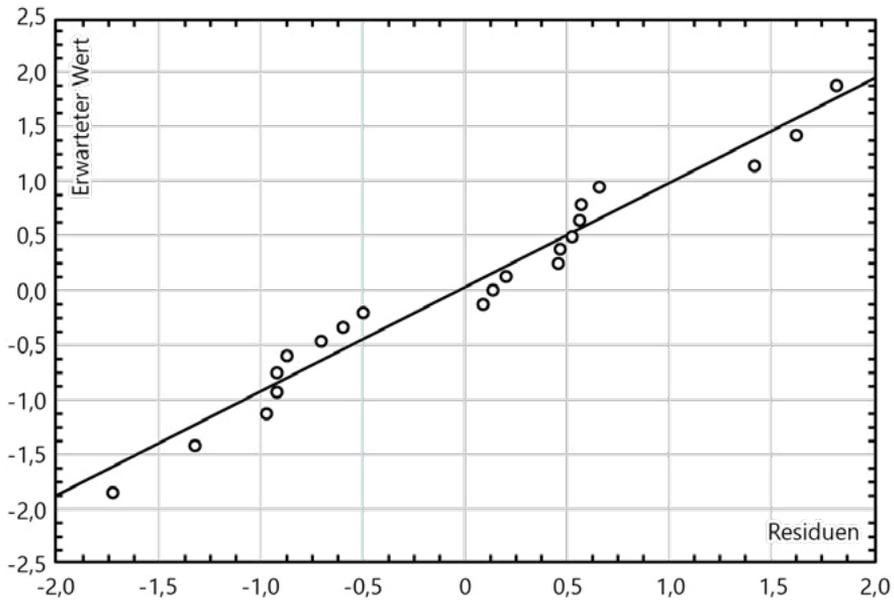


Abb. 6: Normalverteilungs-Plot der Residuen

(6) Die Berechtigung eines linearen Modells des Typs Gleichung (1) wird vorausgesetzt. Dies wird in den Abschnitten 2.10.4 und 2.10.6 für unser Beispiel zu prüfen sein. Wir führen aber hier ein besonders lehrreiches Beispiel aus der Literatur an, dass von ANSCOMBE publiziert wurde. (Man findet es auch bei Wikipedia.) ANSCOMBE hat vier Datenreihen mit jeweils gleichen Mittelwerten der  $X$ - und  $Y$ -Variablen vorgestellt, die bei unbesehen ausgeführter Regressionsrechnung (OLS) alle zum gleichen Resultat führen:  $Y = 3,00 + 0,500 X$  mit einem Bestimmtheitsmaß von  $B = 0,67$  für  $N = 11$  Wertepaare. (Allerdings: Dieses Bestimmtheitsmaß hätte uns sicher alarmiert.)

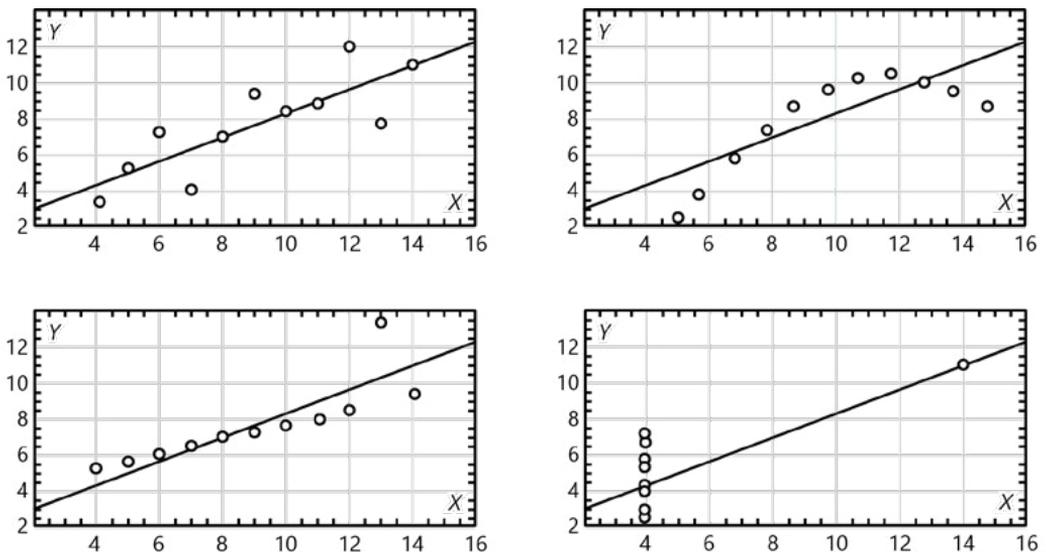


Abb. 7: ANSCOMBE-Quartett

❶ (Daten-ID: F. J. ANSCOMBE: „Graphs in Statistical Analysis“, The American Statistician 27,1(1973)17-22)

### 2.10.3 DATENBEISPIEL ZU 2.2.3 - AUSREIBERTEST

Die Ausreißerfreiheit der Datensätze ist eine entscheidende Voraussetzung für die Zuverlässigkeit und den Erfolg von Modellierungsversuchen. Welchen grundsätzlichen Wert grafische Darstellungen haben, zeigten wir oben anhand des ANSCOMBE-Quartetts.

Das Datenbeispiel von MILLER weist, wie Abbildung 1 zeigt, keine auffälligen Ausreißer auf.

Wir simulieren jetzt Mehrfachmessungen an sieben Messpunkten:

$n_c$	=	1	2	3	4	5	6	7	
$x$	=	0	1	2	3	4	5	6	$cu$
$y$	=	0,05	3,85	10,00	14,4	20,6	26,7	28,7	$su$
		0,10	3,75	10,05	14,3	20,8	27,1	29,1	$su$
		0,15	3,80	9,95	14,5	20,7	26,9	29,5	$su$

Wir prüfen die „verdächtigen“ (nicht optimierten) Werte aus Gruppe  $n_c = 7$ : {28,7 | 29,1 | 29,5} und bilden die Prüfwerte unter Verwendung der unten stehenden Tabelle:

$$d_{MIN} = 0,4/0,8 = 0,5 \text{ und } d_{MAX} = 0,4/0,8 = 0,5$$

Für 5% Irrtumsrisiko und  $N = 3$  beträgt der  $d_{CRIT} = 0,941$ .

Es besteht also kein Grund zur Annahme eines Ausreißers in dieser Kalibriergruppe.

Anzahl $N$ Messwerte	P=95%	P=99%	Minimalwert unterer Prüfwert	Maximalwert oberer Prüfwert
3	0,941	0,988		
4	0,765	0,889		
5	0,642	0,780	$d_{MIN} = \frac{y_2 - y_1}{y_N - y_1}$	$d_{MAX} = \frac{y_N - y_{N-1}}{y_N - y_1}$
6	0,560	0,698		
7	0,507	0,637		
8	0,554	0,683		
9	0,512	0,635	$d_{MIN} = \frac{y_2 - y_1}{y_{N-1} - y_1}$	$d_{MAX} = \frac{y_N - y_{N-1}}{y_N - y_2}$
10	0,477	0,597		
11	0,576	0,679		
12	0,546	0,642	$d_{MIN} = \frac{y_3 - y_1}{y_{N-1} - y_1}$	$d_{MAX} = \frac{y_N - y_{N-2}}{y_N - y_2}$
13	0,521	0,615		
14	0,546	0,641		
15	0,525	0,616		
16	0,507	0,595		
17	0,490	0,577	$d_{MIN} = \frac{y_3 - y_1}{y_{N-2} - y_1}$	$d_{MAX} = \frac{y_N - y_{N-2}}{y_N - y_3}$
18	0,475	0,561		
19	0,462	0,547		
20	0,450	0,535		
21	0,440	0,524		
22	0,430	0,514		
23	0,421	0,505		
24	0,413	0,497		
25	0,406	0,489		
26	0,399	0,482		
27	0,393	0,475		
28	0,387	0,469		
29	0,381	0,463		

Tabelle der kritischen Werte  $d_{CRIT}$   
(nach R. B. DEAN, W. J. DIXON; Analytical Chemistry 23 (1951) 636-639)

## 2.10.4 DATENBEISPIEL ZU 2.3 - STATISTIK DER KOEFFIZIENTEN

In diesem Datenbeispiel zeigen wir die Anwendung statistischer Tests bei der Prüfung auf Signifikanz der Regressionskoeffizienten und ermitteln deren Vertrauensbereich. Dabei benötigt man die Quantile von Verteilungsfunktionen als „kritische Zufallshöchstwerte“. Wir haben in diesem Büchlein darauf verzichtet, entsprechende Tabellen zum Beispiel mit Werten der STUDENTschen  $t$ -Verteilung oder der FISHERSchen  $F$ -Verteilung anzuhängen, da wir davon ausgehen, dass der Leser beim Nachrechnen ohnehin auf Programme zurückgreift, die diese Funktionswerte berechnen. Die brauchbarsten Tabellen nebst anschaulichen Beispielen findet man in: P. H. MÜLLER, P. NEUMANN, R. STORM, (1979).

(1) Als Fehler der Regression und der Koeffizienten erhalten wir:

$$\begin{array}{ll} s_{\text{RES}}^2 = 6,175/5 = 1,235 \text{ su}^2 & \text{bzw.} \quad s_{\text{RES}} = 1,111 \text{ su} \\ s^2(\hat{a}_1) = 1,235/28 \text{ su}^2/\text{cu}^2 & \text{bzw.} \quad s(\hat{a}_1) = 0,210 \text{ su/cu} \\ s^2(\hat{a}_0) = 1,235(1/7 + 3^2/28) = 0,573 \text{ su}^2 & \text{bzw.} \quad s(\hat{a}_0) = 0,757 \text{ su.} \end{array}$$

(2) Die empirischen Testwerte für die Koeffizienten erhalten wir wie folgt:

$$\begin{array}{l} t_{\text{EMP}}(\hat{a}_0) = |0 - (-0,418)|/0,757 = 0,552 \\ t_{\text{EMP}}(\hat{a}_1) = |0 - 5,139|/0,210 = 24,47 \end{array}$$

Sie sind mit dem kritischen Höchstwert  $t_{\text{CRIT}}(a_p) = t_{\text{CRIT}}(5; 0,975) = 2,57$  zu vergleichen.

Es ergibt sich in unserem Beispiel:

für  $\hat{a}_0$ : Die Hypothese eines Achsenabschnitts von Null wird angenommen.

für  $\hat{a}_1$ : Die Hypothese eines Anstiegs von Null wird abgewiesen.

Was den Anstieg betrifft, können wir nun ziemlich sicher sein, dass unser Regressionsmodell überhaupt gerechtfertigt ist, beim Achsenabschnitt haben wir aber ein Dilemma. Hier stehen wir vor der Frage, künftig ohne Achsenabschnitt zu rechnen oder weiter das vollständige Modell zu verwenden. Meistens wird man sich aus sachlichen (Leerwert-Beitrag) und numerischen Gründen, weil damit zum Beispiel die Residuen kleiner ausfallen, für die Beibehaltung des vollständigen linearen Modells entscheiden.

(3) Als Vertrauensbereiche der Koeffizienten erhalten wir:

$$\begin{array}{l} \text{cnf}(\hat{a}_1) = 5,139 \pm 2,57 \cdot 0,210 = 5,139 \pm 0,540 \text{ su/cu,} \\ \text{und, trotz des Resultats aus obigen Tests, auch} \\ \text{cnf}(\hat{a}_0) = -0,418 \pm 1,946 \text{ su.} \end{array}$$

Da das Vertrauensintervall des Achsenabschnitts die Null enthält, können wir aber wieder schließen, dass der wahre Achsenabschnitt wahrscheinlich Null ist.

## 2.10.5 DATENBEISPIEL ZU 2.4 - ALTERNATIVE KALIBRIERMODELLE

Innerhalb des linearen Kalibriermodells hat man zahlreiche Möglichkeiten einer datenadäquaten Schätzung der Modellparameter. Auf einige gehen wir näher ein.

### (2.4.1) Experimentalfehler

Je nach Art der Annahmen oder des Vorwissens über die Experimentalfehler der Größen in den Kalibriermodellen haben wir folgende Spezialfälle unterschieden:

- (A) Annahme: Der Fehler in Y überwiegt.
- (B) Annahme: Der Fehler in X überwiegt.
- (C) Annahme: Die Empfindlichkeit entspricht dem Anstieg.
- (D) Vorwissen: Das Fehlerverhältnis beider Größen ist bekannt.
- (E) Vorwissen: Das Fehlerverhältnis ist gleich Eins.
- (F) Vorwissen: Der Fehler von Y oder von X ist bekannt.

Wir können nun die Koeffizienten einiger Regressionsschätzungen gegenüberstellen:

- (A)  $\hat{a}_1(\text{OLS}) = 143,9/28 = 5,139 \text{ su/cu}$        $\hat{a}_0 = -0,418 \text{ su}$
- (B)  $\hat{a}_1(\text{INV}) = 745,72/143,9 = 5,182 \text{ su/cu}$        $\hat{a}_0 = -0,547 \text{ su}$
- (C)  $\hat{a}_1(\text{RHA}) = (745,72/28)^{1/2} = 5,161 \text{ su/cu}$        $\hat{a}_0 = -0,482 \text{ su}$
- (E)  $\hat{a}_1(\text{ORTH}) = 5,181 \text{ su/cu}$        $\hat{a}_0 = -0,543 \text{ su}$

### (2.4.2) Inverse Kalibration

Die inverse Regression ergibt

$$\hat{a}_1(\text{IRG}) = 0,193 \text{ cu/su} \qquad \hat{a}_0 = 0,105 \text{ cu}$$

Man rechnet leicht nach, dass  $\hat{a}_1(\text{IRG})$  der Kehrwert von  $\hat{a}_1(\text{INV})$  ist und  $\hat{a}_0$  sich mittels der Mittelwerte  $\bar{x}$  und  $\bar{y}$  aus der Geradengleichung ergibt.

### (2.4.3) Gewichtete Regression

Hierfür gibt die Datenlage keinen Anlass.

### (2.4.4) Robuste Regression

Um die *Robuste Regression* zu illustrieren, nehmen wir drei Fälle mit je drei Wertepaaren  $(x_n, y_n)$  an, wobei die Ausreißer unterstrichen sind:

Fall	(AA)			(BB)			(CC)		
$x_n$ :	1	2	3	1	2	3	1	2	3
$y_n$ :	4	6	8	4	<u>9</u>	8	4	6	8

Eine einfache Skizze macht die Problematik deutlich.

Daraus bzw. mittels üblicher Regression berechnen wir:

$$(\text{AA}): \hat{a}_0 = 2, \hat{a}_1 = 2 \quad (\text{BB}): \hat{a}_0 = 3, \hat{a}_1 = 2 \quad (\text{CC}): \hat{a}_0 = 4,857, \hat{a}_1 = 0,857$$

Eine robuste Schätzung im Fall (BB) nach den oben angegebenen Formeln ergibt die drei Koeffizienten:

Paar	1,2	2,3	1,3	
$\hat{a}_1$	5	-1	2	$\text{med}(\hat{a}_1) = 2$
$\hat{a}_0$	-1	11	2	$\text{med}(\hat{a}_0) = 2$

Für die drei Paare von Koeffizienten lassen sich nun auch leicht die Residuen errechnen. Auch ihr Vergleich stützt die Entscheidung für das Modell, das dem ausreißerfreien Fall (AA) entspricht:

$$Y = 2 + 2 \cdot X.$$

Man beachte, dass Ausreißer in  $X$  besonders kritische Situationen darstellen, wie wir in Abschnitt 2.10.3 illustriert haben.

### 2.10.6 DATENBEISPIEL ZU 2.5 - NICHTLINEARE MODELLE

Alternativ zu den linearen Kalibriermodellen gibt es wiederum zahlreiche Möglichkeiten zur datengerechten Modellierung. Wir illustrieren im Folgenden lediglich einen quasilinearen Ansatz.

Gehen wir also versuchsweise der Idee nach, einen nichtlinearen Ansatz zu probieren. Dazu fügen wir der bekannten Wertetabelle einfach die Werte für  $x^2_n$  an:

$n$	=	1	2	3	4	5	6	7	
$x_n$	=	0	1	2	3	4	5	6	$cu$
$x^2_n$	=	0	1	4	9	16	25	36	$cu^2$
$y_n$	=	0,1	3,8	10,0	14,4	20,7	26,9	29,1	$su$

Unter Verwendung einer geeigneten Software<sup>36</sup> erhalten wir die folgenden Ergebnisse:

$$\begin{aligned} \hat{a}_0 &= -0,638 \text{ su} & s(\hat{a}_0) &= 1,070 \text{ su} & t_{\text{EMP}} &= 0,596 & (P \approx 58\% \text{ in } t_{\text{CRIT}}) \\ \hat{a}_1 &= 5,404 \text{ su/cu} & s(\hat{a}_1) &= 0,835 \text{ su/cu} & t_{\text{EMP}} &= 6,468 & (P \approx 0,3\% \text{ in } t_{\text{CRIT}}) \\ \hat{a}_2 &= -0,044 \text{ su/cu}^2 & s(\hat{a}_2) &= 0,134 \text{ su/cu}^2 & t_{\text{EMP}} &= 0,329 & (P \approx 76\% \text{ in } t_{\text{CRIT}}) \end{aligned}$$

Aus den hohen Wahrscheinlichkeitswerten  $P$  können wir ableiten, dass sowohl  $\hat{a}_0$  als auch der Koeffizient  $\hat{a}_2$  des nichtlinearen Terms nur zufällig von Null abweichende Werte haben. Das heißt, wir können auf dem einfachen linearen Modell  $Y = f(X)$  beharren.

### 2.10.7 DATENBEISPIEL ZU 2.6 - BEURTEILUNG VON MODELLEN (EINFACHMESSUNGEN)

(2.6.1) Bestimmtheitsmaß und

(2.6.2) Anpassungsgüte

Eine der wichtigsten Kenngrößen linearer Modelle ist das Bestimmtheitsmaß. Nur für diese Modelle darf das Bestimmtheitsmaß  $B_{XY}$  dem Quadrat des Korrelationskoeffizienten,  $r^2$ , gleichgesetzt werden. Dass der Korrelationskoeffizient eines linearen Zusammenhangs auch bei offensichtlich anderer, insbesondere nichtlinearer Datenstruktur plausible Werte annehmen kann, hat uns das ANSCOMBE-Quartett gelehrt (siehe Abschnitt 2.10.2).

Für einige Modellvarianten seien die Bestimmtheitsmaße  $B$  und die Anpassungsgüte ( $GOF$ ) verglichen und getestet:

Modell	OLS	INV	RHA	ORTH	IRG	
$B$	0,992	1,008	1,000	1,008	0,992	
$F_{\text{EMP}}(B)$	599	- 604	$\infty$	-650	599	$F_{\text{CRIT}}(1, 5; 0,95) = 6,61$
$F_{\text{EXP}}(GOF)$	599	604	602	604	599	$F_{\text{CRIT}}(1, 5; 0,95) = 6,61$

Offenbar erklären alle Modelle die Datenvariation in signifikanter Weise.

36 Wir bevorzugen STATISTICA, StatSoft, Tulsa, hier Version für Windows 95. Mit EXCEL, Microsoft Corporation, lassen sich diese einfachen Beispiele ebenfalls schnell nachrechnen.

Üblicherweise wird man zuerst die OLS-Schätzung vornehmen und ihr Ergebnis testen. Daraus würde sich also keine Notwendigkeit einer alternativen Rechnung ergeben. Interessant ist, dass das RHA-Modell exakt 100% der Datenvariation erklärt, während das OLS-Modell sie ein wenig unterschätzt und das INV- und das ORTH-Modell sie „überschätzen“. Hierin zeigt sich das Phänomen der Überanpassung („overfitting“), auf das in Abschnitt 2.2.2 hingewiesen wurde (Voraussetzung 7).

### 2.10.8 DATENBEISPIEL ZU 2.6 - BEURTEILUNG VON MODELLLEN (MEHRFACHMESSUNGEN)

Je mehr Daten man hat, umso verlässlicher (schärfer) kann man die Modellgüte testen. Schon die Erörterungen zur Ausreißerproblematik legen Wiederholungen von Versuchsansätzen, also auch Mehrfachmessungen nahe. Außerdem sehen wir einen Einfluss auf die Darstellungen des Kalibriergraphen.

Im Teil (1) lernen wir die wichtigsten Parameter-Tests kennen. In Teil (2) besprechen wir Tests auf Varianzgleichheit im Kalibrierbereich, eine entscheidende Voraussetzung linearer Modelle.

Durch die Mehrfachmessungen ist es zunächst möglich, den linearen Kalibriergraphen durch Fehlerbalken weiter zu charakterisieren. Man kann zugleich feststellen, dass das Konfidenzband schmaler wird.

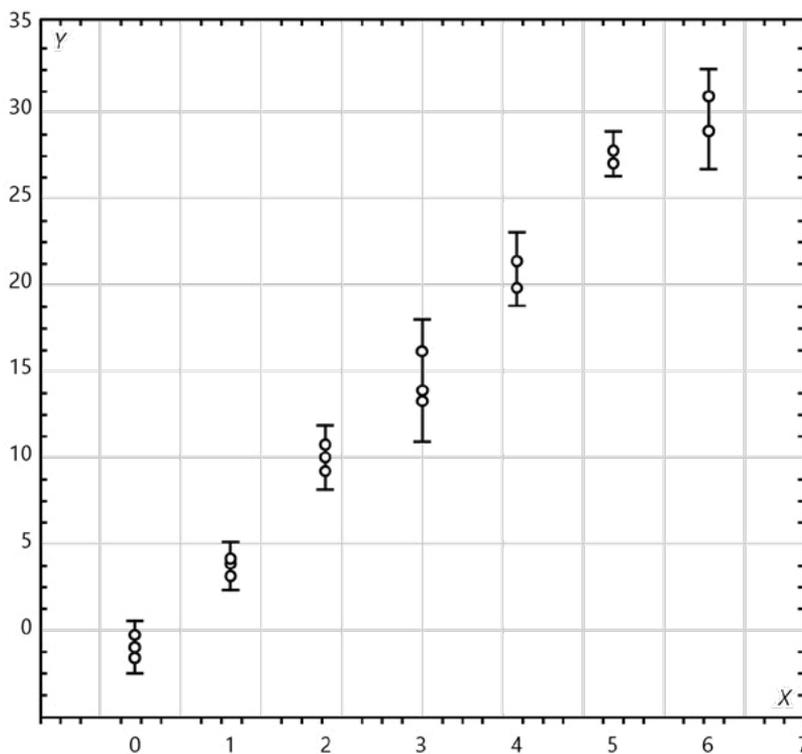


Abb. 8: Signalwerte mit Fehlerbalken

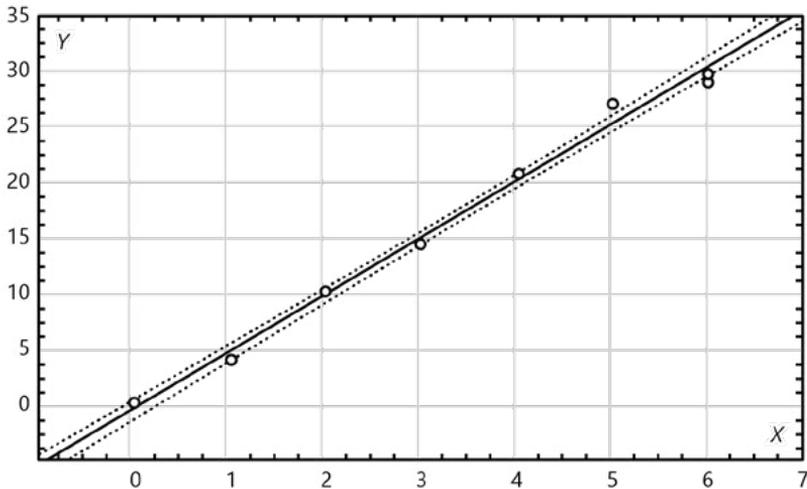


Abb. 9: Kalibriergraph mit Konfidenzband

### (2.6.3) Anpassungsdefekt

(1) Um den Anpassungsdefekt (*LOF*) und die Anpassungsgüte (*TOA*) testen zu können, erweitern wir unser Datenbeispiel (nach MILLER, 1991) durch Simulation von  $W_{\text{REP}} = 3$  Wiederholungen der Signalmessung an jedem der  $N_c = 7$  Kalibrierpunkte. Die  $N = 3 \cdot 7 = 21$  *y*-Werte sind nachfolgend angegeben.

(Wir haben es so eingerichtet, dass die jeweiligen Mittelwerte der Wiederholungen den Originalmesswerten entsprechen.)

$n_c$	=	1	2	3	4	5	6	7	
$x$	=	0	1	2	3	4	5	6	<i>cu</i>
$y$	=	0,05	3,85	10,00	14,4	20,6	26,7	28,7	<i>su</i>
		0,10	3,75	10,05	14,3	20,8	27,1	29,1	<i>su</i>
		0,15	3,80	9,95	14,5	20,7	26,9	29,5	<i>su</i>

$$F_{\text{EMP}}(\text{GOF}) = (2218,6/1) / (18,98/19) = 2221 \quad \text{vgl. mit } F_{\text{CRIT}}(1, 19; 0,95) = 4,38$$

$$F_{\text{EMP}}(\text{LOF}) = (18,53/5) / (0,451/14) = 115 \quad \text{vgl. mit } F_{\text{CRIT}}(5, 14; 0,95) = 2,96$$

$$F_{\text{EMP}}(\text{TOA}) = (18,98/19) / (0,451/14) = 31 \quad \text{vgl. mit } F_{\text{CRIT}}(19, 14; 0,95) = 2,40$$

Wie bei der vorangegangenen Analyse zeigt der *GOF*-Test an, dass das Regressionsmodell zur Beschreibung der Daten geeignet ist.

Der *LOF*-Test bestätigt nun, dass hochpräzise Messungen vorliegen und damit der Modellfehler größer als der Experimentalfehler ist. Wie auch auf der Basis des *TOA*-Tests wird man sich möglicherweise nun dafür entscheiden, ein anderes, vielleicht nichtlineares, Modell zu probieren.

(2) Zur Prüfung auf Varianzgleichheit im Kalibrierbereich ergänzen wir die obigen Daten um die Zeile der Varianzen an den Messpunkten:

$n_c$	=	1	2	3	4	5	6	7	
<i>var</i>	=	0,0025	0,0025	0,0025	0,01	0,01	0,04	0,16	$su^2 \Sigma = 0,2275$
<i>s</i>	=	0,0500	0,0500	0,0500	0,10	0,10	0,20	0,40	<i>su</i>

Zur Prüfung auf Varianzgleichheit (Homoskedastizität) kann man zwei Tests heranziehen, die insofern Verallgemeinerungen des bekannten  $F$ -Tests sind, als sie die Hypothese der Gleichheit von mehr als zwei Varianzen prüfen. Trotz des von der ISO<sup>37</sup> empfohlenen COCHRAN-Tests wird von manchen Autoren am BARTLETT-Test festgehalten.<sup>38</sup>

Gegen den BARTLETT-Test spricht, dass er wesentlich empfindlicher bei Abweichungen von der Normalverteilung ist. Auch ist der BARTLETT-Test nicht anwendbar, wenn eine Varianz gleich Null ist. Vorteilhaft ist jedoch, dass man gegen kritische Werte der  $\chi^2$ -Verteilung testet, die leicht zugänglich ist.

Beim BARTLETT-Test erhalten wir in unseren Fall einen Testwert von  $\chi^2_{EMP} = 16,83$ , der mit  $\chi^2(6; 1 - 0,05) = 12,59$  zu vergleichen ist. Damit würde die Hypothese der Gleichheit der sieben Varianzen abgelehnt. Es ergibt sich sofort die Frage, *an welcher Stelle* die Varianz wohl unakzeptabel ist. Natürlich können wir aus Betrachtung der oben gegebenen Werte sofort eine Vermutung aufstellen.

Gehen wir aber nun korrekt nach dem COCHRAN-Test vor.

Dieser beruht auf einer recht einfachen Testgröße, erfordert allerdings die gleiche Anzahl von Wiederholungen  $W_{REP}$  für jede Kalibrierstelle. Leider ist die Vergleichsgröße  $C$  selten tabelliert. Man findet die Tabellen zum Beispiel im oben zitierten ISO Standards Handbook von 1989 oder bei MÜLLER et al. (1979).

Die entsprechenden Ausdrücke sind:

$$C_{EMP} = \frac{\max s_n^2}{\sum_{n=1}^{N_C} s_n^2} \quad \text{vs.} \quad C_{CRIT} = C(N_C, W_{REP}-1; \alpha) \quad (59)$$

Beim COCHRAN-Test erhalten wir  $C_{EMP} = 0,16 / 0,2275 = 0,703$ .

Da der Wert größer als  $C_{CRIT} = C(7, 2; 0,05) = 0,5612$  ausfällt, wird die Hypothese der Varianzgleichheit abgelehnt und gleichzeitig die Varianz der Messwertgruppe 7 als zu hoch erkannt.

Schon die Betrachtung der Rohdaten zeigt eine typische Varianzsituation, für die wahrscheinlich eine gewichtete Regression der Ausweg wäre.

Es gibt aber zwei andere Möglichkeiten, zu reagieren.

(1) Erstens könnte man, wenn es sachlich erlaubt ist, auf die Modellierung des weitgespannten Kalibrierbereichs verzichten und die *Daten der Gruppe 7 ignorieren*.

(2) Zweitens könnte man die Experimente, wenigstens die der Gruppe 7, unter optimierten Bedingungen wiederholen.

37 vgl. ISO 5725-1986, in ISO Standards Handbook 3, S. 427f. und Tabelle mit kritischen Werten S. 458

38 U. HILLEBRAND: „Der F-Test und die Kalibrierfunktion“, GIT Labor-Fachzeitschrift Heft 8(1999)820-822, U. HILLEBRAND: „Varianzhomogenität bei der Kalibration von Analysenverfahren“, LaborPraxis Februar (2000)64-65

(1) Auf Daten verzichten

Nach Ignorieren der Messwertgruppe 7 erhalten wir  $C_{EMP} = 0,04 / 0,0675 = 0,5926$ . Da dieser Wert kleiner als  $C_{CRIT} = 0,6161$  ausfällt, wird die Hypothese der Varianzgleichheit nunmehr angenommen.

Natürlich muss man nun die Regressionsparameter neu berechnen:

$$\begin{array}{ll} \hat{a}_1 (\text{o. Gruppe 7}) = 5,403 \text{ su/cu} & \hat{a}_0 = -0,857 \text{ su} \\ s(\hat{a}_1) = \pm 0,103 \text{ su/cu} & s(\hat{a}_0) = \pm 0,312 \text{ su} \quad s_{RES} = 0,747 \text{ su} \end{array}$$

Wir erinnern zum Vergleich an:

$$\begin{array}{ll} \hat{a}_1 (\text{OLS/3 Wh.}) = 5,139 \text{ su/cu} & \hat{a}_0 = -0,418 \text{ su} \\ s(\hat{a}_1) = \pm 0,109 \text{ su/cu} & s(\hat{a}_0) = \pm 0,393 \text{ su} \quad s_{RES} = 1,000 \text{ su} \end{array}$$

Kein Wunder, dass die Koeffizienten bei OLS/3 Wh. mit den OLS-Werten übereinstimmen, hatten wir doch dafür gesorgt, dass die jeweiligen Signalmittelwerte mit den früheren Einzelwerten übereinstimmten. Beim Restfehler der Regression macht sich allerdings der größere Freiheitsgrad bemerkbar.

(2) Daten optimieren

Nach der (optimierten) Wiederholung an der Kalibrierstelle 7 sollen sich folgende Signalwerte ergeben haben: {28,9 | 29,1 | 29,3}.

(Wir haben damit die gleichen absoluten Differenzen wie in Gruppe 6.)

Wir erhalten jetzt  $C_{EMP} = 0,3721$ . Da der Wert kleiner als  $C_{CRIT} = 0,5612$  ausfällt, wird die Hypothese der Varianzgleichheit wiederum angenommen.

Die neu berechneten Regressionsparameter sind:

$$\begin{array}{ll} \hat{a}_1 (\text{Gruppe 7 opt.}) = 5,139 \text{ su/cu} & \hat{a}_0 = -0,418 \text{ su} \\ s(\hat{a}_1) = \pm 0,108 \text{ su/cu} & s(\hat{a}_0) = \pm 0,391 \text{ su} \quad s_{RES} = 0,993 \text{ su} \end{array}$$

## 2.10.9 DATENBEISPIEL ZU 2.7 - SIGNALBEZOGENE STATISTIK (VARIANZANALYSE)

Mit der Prüfung auf Varianzgleichheit im vorigen Abschnitt näherten wir uns bereits einem mächtigen datenanalytischen Werkzeug, der statistischen Varianzanalyse. Sie untersucht im Detail die Auswirkung von Einflussgrößen (als Faktoren) auf Messgrößenwerte. Wir wollen hier nicht annähernd ihre methodischen Feinheiten ausloten, sondern nur darauf hinweisen, dass sie sich ausschließlich auf die gemessenen Werte bezieht, während die Einflussgrößen als Faktoren wirken, die diese Messwerte in Datengruppen (Klassen) gliedern.

Im Teil (1) illustrieren wir das Grundprinzip anhand der üblichen Auswertetabelle für eine einfaktorielle Varianzanalyse.

In Teil (2) wird ein Auswertetableau mit Quadratesummen (SAQ) angegeben, mit dem auf regressionsanalytische Feinheiten eingegangen werden könnte. Die mittleren Abweichungsquadrate (MAQ) ergeben sich jeweils als Quotient von SAQ und zugehörigem Freiheitsgrad.

Möge der Leser selbst mittels seiner bevorzugten Software anhand der obigen Zahlenwerte (Werte der Gruppe 7 nicht „optimiert“) eine einfaktorielle *Varianzanalyse* ausführen.

Dabei wäre zu überprüfen, ob die damit erhaltenen ANOVA-Tabellen die folgenden Quadratsummen bzw. Freiheitsgrade enthalten und ob die Additivitätsbedingungen erfüllt sind.

(1) Auswertetabelle zur einfachen Varianzanalyse

In die im Abschnitt 2.7 gezeigte Tabelle können wir nun konkrete Werte eintragen:

Varianzursache	SAQ	fg	MAQ	$F_{EMP}$
zwischen X-Stufen	2237,16	6	372,860	11473
innerhalb X-Stufen	0,46	14	0,033	
total	2237,62	20		

Die Additivität der SAQ und der Freiheitsgrade trifft zu.

Der Testwert  $F_{EMP}$  ist gegen  $F_{CRIT}(6, 14; 0,99) = 4,46$  zu testen.

Da der kritische Höchstwert für eine Normalverteilung der Gruppenmittel überschritten wird, ist das signifikante Wirken des Konzentrationsfaktors bestätigt.

(2) Verfeinerte Auswertetabelle

In ähnlicher Weise kann die umfangreichere Auswertetabelle gefüllt werden, deren Inhalt wir zunächst in allgemeiner Form angeben. Dabei sind einige der in der Literatur gebräuchlichen Bezeichnungen der SAQ berücksichtigt.

	Summenterm =	Summand 1 +	Summand 2
<b>SAQ</b>	$SS_{TOT} - SS'_{YY}$	$SS_{CENTr} - SS_{YY}$	$SS_{MEAN}$
<b>fg</b>	$N$	$N - 1$	$1$
<b>SAQ</b>	$SS_{CENTr} - SS_{YY}$	$SS_{RES}$	$SS_{FACTr} - SS_{MOD}$
<b>fg</b>	$N - 1$	$N - P$	$P - 1$
<b>SAQ</b>	$SS_{RES}$	$SS_{LOF}$	$SS_{EXP}$
<b>fg</b>	$N - P$	$N_{REP} - P$	$N - N_{REP}$

Der Wert für  $SS'_{YY}$  ergibt sich aus  $SS_{YYr}$ , wenn  $\bar{y} = 0$  ist;  $SS_{MEAN} = N \cdot \bar{y}^2$ .

Mit konkreten Werten gefüllt, stellt sich das Tableau so dar:

	Summenterm =	Summand 1 +	Summand 2
<b>SAQ</b>	6963	2238	4725
<b>fg</b>	21	20	1
<b>SAQ</b>	2238	18,99	2219
<b>fg</b>	20	19	1
<b>SAQ</b>	18,99	18,53	0,46
<b>fg</b>	19	5	14

Der Leser mag die Additivität selbst prüfen.

Mit einigen Werten lassen sich alle bisher benutzten Testgrößen berechnen, wenn man die entsprechenden MAQ als Zähler oder Nenner verwendet.

	Summenterm =	Summand 1 +	Summand 2
<b>SAQ</b>	$SS_{TOT}, SS'_{YY}$	$SS_{CENT}, SS_{YY}$	$SS_{MEAN}$
<b>fg</b>	$N$	$N - 1$	
<b>SAQ</b>	$SS_{CENT}, SS_{YY}$	$SS_{RES}$	$SS_{FACT}, SS_{MOD}$
<b>fg</b>	$N - 1$	$N - P$	$P - 1$
	Nummer B	Nummer GOF	Zähler B, Zähler GOF
<b>SAQ</b>	$SS_{RES}$	$SS_{LOF}$	$SS_{EXP}$
<b>fg</b>	$N - P$	$N_{REP} - P$	$N - N_{REP}$
	Zähler TOA	Zähler LOF	Nummer LOF, Nummer TOA

## 2.10.10 DATENBEISPIEL ZU 2.7 - ANALYTBEZOGENE STATISTIK (VERTRAUENSBEREICH)

Ziel quantitativer analytischer Bestimmungen ist es, richtige und möglichst präzise Werte eines oder mehrerer Analyte zu ermitteln. Infolge möglicher zufälliger Einflüsse wird man jedoch niemals punktgenaue Ergebnisse erhalten. Vielmehr wird man von einem Vertrauens- oder Konfidenzbereich um jeglichen Analysenwert ausgehen müssen.

(1) Sind Einzelwerte ausreichend?

Um ein zuverlässiges Ergebnis für Analysenwerte zu erhalten, sind in der Regel Mehrfachmessungen anzuraten, woraus dann der Mittelwert gebildet werden kann.

In Ausnahmefällen muss man sich mit Einzelmessungen begnügen.

MILLER hat dafür ein Testkriterium angegeben, bei dessen Erfüllung,  $g < 0,05$ , man im Anwendungsschritt mit Einzelergebnissen vorlieb nehmen darf.

Das MILLERSche Testkriterium lautet  $g = t^2(df; q) s^2_{RES} / (a^2_1 SS_{XX})$ .

Zu seiner Berechnung verwenden wir folgende Werte:

$$s^2_{RES} = 1,235$$

$$a_1 = 5,139$$

$$SS_{XX} = 28$$

$$t^2(5; 0,975) = 6,608$$

wobei

$$df = N_c - 2 \text{ und } \alpha = 0,5 \text{ (!) in } t(df; q) = 2,571.$$

Damit erhalten wir  $g = 0,011$ , weshalb wir mit  $N_A = 1$  verfahren dürfen.

(2) Messwerte für Analysenproben

Als fiktives Ergebnis für zwei Analysenproben, A1 und A2, mögen wir

$$y_A(A1) = 16 \text{ su und } y_A(A2) = 27 \text{ su}$$

erhalten haben.

(Mit der ersten Testprobe simulieren wir ein Signal nahe dem Datenschwerpunkt, mit der zweiten Testprobe nähern wir uns der oberen Grenze der Kalibrierdomäne.)

(3) Rechenwerte für Analysenergebnisse

Aus dem OLS-Modell ( $a_1 = 5,139$  und  $a_0 = -0,418$ ) ergeben sich direkt die Analysenwerte

$$x_A(A1) = 3,195 \text{ und}$$

$$x_A(A2) = 5,335.$$

(4) Vertrauensbereich für Analysenergebnisse

Aus Gleichung (52) ergeben sich somit folgende Vertrauensbereiche  $cnf$  für die Analysenwerte:

$$A1: \quad cnf(x_A) = 3,195 \pm 0,595$$

$$A2: \quad cnf(x_A) = 5,335 \pm 0,643.$$

Man erkennt gut, wie sich das an den beiden Enden des Kalibrierbereichs aufgeweitete Konfidenzband um den Kalibriergraphen auswirkt.

## 2.10.11 DATENBEISPIEL ZU 2.9 - VERGLEICH VON REGRESSIONSGERADEN

Der Vergleich von Regressionsgeraden ist bei verschiedenen Sachverhalten erforderlich, so etwa beim Vergleich unterschiedlicher Analysenverfahren oder Messmethoden für den gleichen Analyten innerhalb eines Labors. Das Thema greifen wir in Abschnitt B auf. Auch bei der Qualitätssicherung mittels Methodenvergleich innerhalb von Ringversuchen zwischen verschiedenen Labors kann er bedeutsam sein. Darauf gehen wir gleich kurz ein. (Der Leser könnte versuchen, zum Beispiel mit den bei STORM (1995) angegebenen Formeln, die folgenden Ergebnisse nachzuvollziehen.)

Wir beziehen uns auf bereits erhaltene Ergebnisse und deklarieren die OLS-Ergebnisse aus den Wiederholungsmessungen an allen 7 Messstellen als von **Labor A** stammend und die Ergebnisse aus dem verkürzten Kalibrierbereich mit nur 6 Messstellen als zu **Labor B** gehörend.

Zur Berechnung brauchen wir folgende Werte:

	Labor A	Labor B
	(mit streuender/„optimierter“ Messwertgruppe 7)	
$N_C =$	7	6
$N_{REP} =$	7	6
$W_{REP} =$	3	3
$N =$	21	18
$fg_{CAL} =$	19	16
$\bar{x}$	3	2,5
$\bar{y}$	15	12,65
$SS_{XX} =$	84	52,5
$SS_{RES} =$	18,985	8,940
$a_1 =$	5,139	5,403
$a_0 =$	-0,418	-0,857

(1) Prüfen der „Reststreuung“ der beiden Modelle

Der Vergleich von  $F_{EMP} = 0,999/0,559 = 1,787$  mit  $F_{CRIT}(19, 16; 0,975) = 2,70$  zeigt, dass die Modelle vergleichbar sind. Der übliche „doppelte“  $t$ -Test ist erlaubt.

(2) Prüfen der Anstiege

Der Vergleich von  $t_{EMP} = |-0,264| / 0,157 = 1,68$  mit  $t_{CRIT}(35; 0,975) = 2,03$  weist auch die Anstiege als gleich aus.

(3) Prüfen der Konstanten

Der Vergleich von  $t_{EMP} = 0,934$  mit  $t_{CRIT}(35; 0,975) = 2,03$  ergibt ebenfalls keinen Hinweis auf Unterschiedlichkeit der Regressionsgeraden.

Es kann also angenommen werden, dass beide Regressionsgeraden nur zufällig voneinander abweichen.

## 3 Nachweisgrenzen und andere Grenzwerte

---

### 3.1 Begriffliche Probleme

Eine Information über die Leistungsfähigkeit einer Analysenmethode bzw. eines Analysenverfahrens erhält man aus Präzisionsmaßen (Messpräzision und Methodenpräzision) und Angaben zur „Richtigkeit“ (Ergebnisunsicherheit).

Die *Messpräzision* ist ein Maß für die Unsicherheit der Messwerte (Signal-domäne) und wird durch wiederholtes „Messen“ eines „Standards“ ermittelt.

Die *Methodenpräzision* ist ein Maß für die Unsicherheit des Analyseergebnisses (zum Beispiel in der Konzentrationsdomäne) und bezieht sich auf die mehrfache Durchführung einer Analyse. Umfassender ist vielleicht der Begriff *Verfahrenspräzision*.

Hier kann, vor allem im Kontext von Ringversuchen, noch zwischen Wiederholpräzision (Reproduzierbarkeit) und Vergleichspräzision unterschieden werden. Auf die Problematik dieser Maßzahlen gehen wir nicht weiter ein.

Abgesehen von den systematischen und zufälligen Verfälschungen der Messwerte in den einzelnen Verfahrensschritten trägt bei den kalibrierbedürftigen Verfahren außerdem noch der Kalibriervorgang zur Unsicherheit der Analyseergebnisse bei. Ein vollständig definiertes Verfahren ist deshalb durch den Konzentrationsbereich zu beschreiben, der durch die Kalibrierfunktion zugänglich ist, und durch die Kenngrößen der Kalibrierfunktion selbst.

Aus zwei Kenngrößen der Kalibrierfunktion wird häufig die *Verfahrensstandardabweichung* berechnet, die ihren Namen eigentlich nur verdient, wenn die Messwerte nach Durchlaufen des vollständigen Verfahrens gewonnen worden sind. Mit der Modellierung der Kalibrierfunktion beschäftigte sich der vorige Abschnitt.

Schließlich spielen bei der Auswahl bzw. Bewertung eines Analysenverfahrens auch noch *Grenzwerte* eine Rolle, die aus oder mittels Kalibrierfunktionen gefunden werden können.<sup>39</sup>

So sind Nachweisgrenzen vor allem bei spurenanalytischen Verfahren interessant. Sie werden manchmal auch zum Vergleich von Analysenmethoden herangezogen, sollten aber nur zur Charakterisierung von Verfahren verwendet werden.

Die Nachweisgrenze ist ein wichtiges Kriterium bei der Optimierung eines spurenanalytischen Verfahrens. Sie hängt stark von möglichen Fehlern in den einzelnen Verfahrensschritten ab.

---

39 H. MÜLLER, T. BRAUN, E. SORKAU: „Der Einfluss des gewählten chemometrischen Modells auf berechnete Nachweisgrenzen in der GF AAS“ in B. WELZ (Hrsg.): CANAS '95 Colloquium Analytische Atom-spektroskopie, Bodenseewerk Perkin Elmer, Überlingen, 1996, 237-243

Zu ihrer Ermittlung gibt es zahlreiche Varianten, die mehr oder weniger kompliziert erscheinen<sup>40</sup> und an deren Verbindlichkeit ständig gearbeitet wird.

Die genaue Angabe des Berechnungsmodus dieser Verfahrenskenngröße sollte deshalb unbedingt Bestandteil einer vollständigen Analysenvorschrift sein.

Der Gebrauch des Begriffs Nachweisgrenze und anderer Grenzwertbegriffe ist in der Literatur nicht einheitlich und bedarf dringend der internationalen Harmonisierung (vgl. Abschnitt 3.4).

## 3.2 Besonderheiten in der Spurenanalytik

In der Spurenanalytik ist insbesondere mit gravierenden Auswirkungen systematischer Fehler zu rechnen, die zu einseitigen Ergebnisverschiebungen führen. Dies ist vor allem bei der mobilen Analytik im Umweltbereich zu beachten.<sup>41</sup>

Man kann grob zwei Fälle systematischer Signalverfälschung unterscheiden.

### Fall A:

Bei Abwesenheit ( $x_A = 0$ ) des zu bestimmenden Spurenbestandteils in einer Probe kann dessen Vorhandensein vorgetäuscht werden, zum Beispiel, wenn der Messwert  $y_n$  den kritischen Wert (level)  $y_{\text{CRIT}}$  einer analytischspezifischen Signalgröße übersteigt.

### Fall B:

Trotz Vorhandenseins des Analyten ( $x_A > 0$ ) liegt der Messwert unter dem kritischen Wert,  $y_n < y_{\text{CRIT}}$ . Der Fall, dass ein Analytischsignal entsteht, das nicht mehr dem Spurenbereich zugeordnet werden kann,  $y_n \gg y_{\text{CRIT}}$ , ist hier ohne Bedeutung.

Das Risiko einer Fehlentscheidung bei der Signalbewertung kann in beiden Fällen durchaus unterschiedliches Gewicht haben.

Aus Sicht der Signalverteilung der Leerwerte handelt es sich statistisch bei Fall A um eine Fehlentscheidung 1. Art, nämlich die fälschliche Abweisung eines Leerwertsignals, und bei Fall B um eine Fehlentscheidung 2. Art, nämlich die fälschliche Annahme eines Leerwertsignals.<sup>42</sup>

Diese beiden Fälle, die schließlich zu einer einseitigen Verschiebung des Ergebniswertes vom „wahren“ Wert führen, werden an dieser Stelle nicht weiter diskutiert, da systematische Fehlerquellen prinzipiell erkannt und beseitigt werden können.

---

40 M. LUTHARDT, E. THAN, H. HECKENDORFF: „Nachweis-, Erfassungs- und Bestimmungsgrenze analytischer Verfahren“, Fresenius J. Anal. Chem. 326(1987)331-339. DIN 32 645 (Mai 1994): „Chemische Analytik; Nachweis-, Erfassungs- und Bestimmungsgrenze; Ermittlung unter Wiederholbedingungen; Begriffe, Verfahren, Auswertung“, Beuth Verlag, Berlin, Köln

41 H. MÜLLER, H. W. ZWANZIGER, J. FLACHOWSKY: „Trace Analysis“, Ullmann's Encyclopedia of Industrial Chemistry, Vol. B5, Wiley-VCH, Weinheim, 2001

42 H. ZWANZIGER, U. ROHLAND, G. WERNER: „Risiko 2. Art und Stichprobenumfang bei statistischen Tests“, Fresenius Z. Anal. Chem. 323(1986)371-374

Die unvermeidbaren zufälligen Fehler bewirken eine beidseitige, zufällige Abweichung einzelner Messwerte bzw. Ergebniswerte von einem bestimmten Wert (Sollwert, Mittelwert, Median) und bedingen ebenfalls unterschiedliche Risiken bei der Signalbeurteilung bzw. der Angabe von Analyseergebnissen in der Spurenanalytik.

### 3.3 Der kritische Signalwert und die Verfahrensgrenzen

Bei der Ermittlung der Verfahrenskenngrößen Nachweis-, Erfassungs- und Bestimmungsgrenze kommt dem kritischen Signalwert  $y_{\text{CRIT}}$  eine zentrale Bedeutung zu. Dieser Signalwert führt über die berechnete Kalibrierbeziehung, zum Beispiel nach Gleichung (1), zur Nachweisgrenze  $x_{\text{CRIT}} = x_{\text{DECISION}}$  (siehe folgende Abbildung).

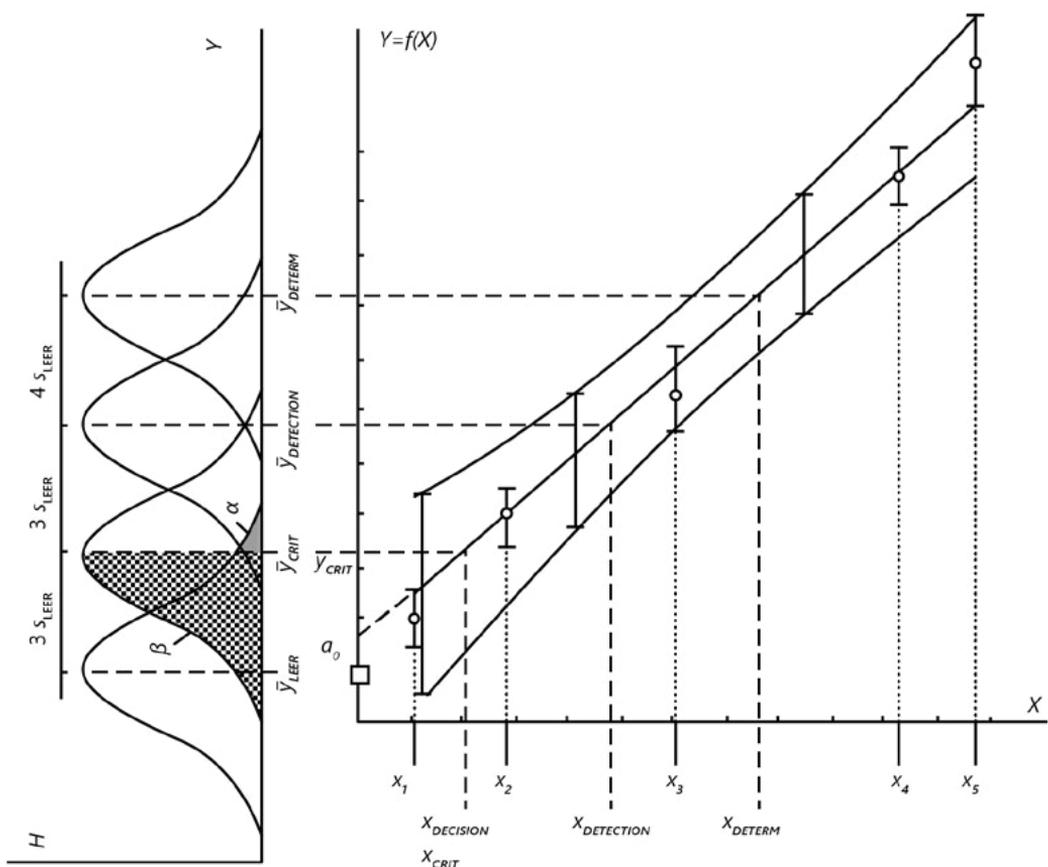


Abb. 10: Illustration von Grenzwerten und Entscheidungsrisiken anhand der Häufigkeitsverteilungen (H) definierter Signalwerte

Am einfachsten ermittelt man den kritischen Signalwert aus wiederholten Analysen mehrerer Leerproben oder, was jedoch weniger reale Kenngrößen liefert, aus wiederholten Messungen einer Leerprobe.

Auch bei nicht vorhandenem Analyten werden dann zum Beispiel insgesamt  $N_L$  zufallsbedingte Signalwerte  $y_l$  ( $l = 1, \dots, N_L$ ) gemessen, die nicht Null sein müssen, und deren Verteilungsparameter leicht berechnet werden können.

$$\begin{aligned}
 \text{Mittelwert} \quad \bar{y}_{LEER} &= \frac{1}{N_L} \sum_{l=1}^{N_L} y_l \\
 \text{Varianz} \quad s_{LEER}^2 &= \frac{1}{N_L - 1} \sum_{l=1}^{N_L} (y_l - \bar{y}_{LEER})^2 \\
 \text{Standard-} \\
 \text{abweichung} \quad s_{LEER} &= \sqrt{s_{LEER}^2}
 \end{aligned} \tag{60}$$

Die obere Grenze des Vertrauensintervalls des Mittelwertes des Leerwertes ist dann als **kritischer Wert**  $y_{CRIT}$  der Messgröße anzusehen. Er berechnet sich nach

$$y_{CRIT} = \bar{y}_{LEER} + t(fg; q) s_{LEER} \sqrt{\frac{1}{N_L}} \tag{61}$$

mittels eines Tabellenwertes für die  $t$ -Verteilung mit  $fg = N_L - 1$  Freiheitsgraden und der Ordnung  $q = 1 - \alpha$ , d. h. mittels einseitiger Quantile.

Es ist hier darauf hinzuweisen, dass man damit von einer symmetrischen Verteilung der Messwertsignale ausgeht. Diese Vorstellung und insbesondere die von normalverteilten Werten ist aber bei Spurenbestimmungen häufig nicht gerechtfertigt.

Evtl. muss eine logarithmische Transformation vorgenommen werden.<sup>43</sup>

Liegt ein Signalwert irgendeiner Messprobe oberhalb dieses kritischen Wertes, so kann mit einem bestimmten Irrtumsrisiko  $\alpha$ , das gewöhnlich auf 5% ( $\alpha = 0,05$ ) festgelegt wird, davon ausgegangen werden, dass kein Leerwertsignal sondern ein analytabhängiges Signal vorliegt.

Hier fällt also seitens des Signals die *qualitative Entscheidung* (engl.: decision) über das Vorhandensein eines Analyten.

Die entsprechende kritische Konzentration  $x_{DECISION} = x_{CRIT}$ , die *Nachweisgrenze*, erhält man durch Verknüpfen der Kalibrierbeziehung mit dem Ausdruck für  $y_{CRIT}$  unter der Annahme, dass  $\bar{y}_{LEER} = a_0$ , als

$$x_{DECISION} = \frac{t(fg; q) s_{LEER}}{a_1} \sqrt{\frac{1}{N_L}} \tag{62}$$

43 Siehe KAISER (1973); K. DOERFFEL, G. MICHAELIS: „Auswertung eines Ringversuchs im Spurenbereich“, Fresenius Z. Anal. Chem. 328(1987)226-227

Daraus ergeben sich unmittelbar als Möglichkeiten zur Optimierung der Nachweisgrenze: Verringern der Leerwertstandardabweichung, Erhöhen der Empfindlichkeit  $a_1$ , Erhöhen der Anzahl  $N_L$  der Leerwertmessungen.

Mit dem Quantil der Normalverteilung  $u(q = 0,999) = 3$  anstelle des  $t$ -Wertes wird daraus die Definition von KAISER (1965).

Der numerische Wert der Nachweisgrenze kann auch durch Berücksichtigen des Aufwandes im zukünftigen Analysenschritt, d. h. die Anzahl  $N_A$  der Messungen der Analysenprobe, verbessert werden. Nimmt man für die eigentliche Analyse die gleiche Standardabweichung wie für die Leerwertmessungen an, so ergibt sich aus dem kritischen Signal  $y_{CRIT}$  die Nachweisgrenze  $x_{DECISION}$

$$y_{CRIT} = \bar{y}_{LEER} + t(fg; q) s_{LEER} \sqrt{\frac{1}{N_L} + \frac{1}{N_A}} \quad (63)$$

$$x_{DECISION} = \frac{t(fg; q) s_{LEER}}{a_1} \sqrt{\frac{1}{N_L} + \frac{1}{N_A}} \quad (64)$$

mit  $fg = N_L - 1$  und  $q = 1 - \alpha$  (einseitig!).

Ein zweiter Zugang zum kritischen Signalwert besteht in der Verwendung der oberen Vertrauensgrenze des Achsenabschnittes  $a_0$  der Kalibrierbeziehung, die aus  $N_C$  Messungen von Kalibrierproben abgeleitet wurde (DIN-Variante). Durch Einsetzen dieses Ausdrucks in das Kalibriermodell ergibt sich wieder die Nachweisgrenze.

$$y_{CRIT} = a_0 + t(fg; q) s_{RES} \sqrt{\frac{1}{N_C} + \frac{1}{N_A} + \frac{\bar{x}_C^2}{SS_{XX}}} \quad (65)$$

$$x_{DECISION} = \frac{t(fg; q) s_{RES}}{a_1} \sqrt{\frac{1}{N_C} + \frac{1}{N_A} + \frac{\bar{x}_C^2}{SS_{XX}}} \quad (66)$$

mit  $fg = N_C - 2$  und  $q = 1 - \alpha$  (einseitig!).

Die Reststandardabweichung  $s_{RES}$  des Kalibriermodells und die Summe der Abweichungsquadrate  $SS_{XX}$  wurden im Abschnitt 2.3.1 definiert.

Die Nachweisgrenze könnte nun weiter dadurch verbessert werden, dass der Mittelwert des Kalibrierbereiches möglichst nahe an der Nachweisgrenze liegt.

Aus Abbildung 10 ist zu erkennen, dass die statistische Verteilung von Messwerten an der Nachweisgrenze  $x_{\text{CRIT}}$  mit  $\beta = 50\%$  Wahrscheinlichkeit Werte liefert, die in den Streubereich des mittleren Leerwertes fallen; siehe Verteilung um  $\bar{y}_{\text{CRIT}}$ .

Ein Analytsignal ist andererseits dann statistisch *sicher erfasst*, wenn das Risiko möglichst gering ist, dass das Signal kleiner als das Nachweisgrenzen-Signal  $\bar{y}_{\text{CRIT}}$  ist; siehe Verteilung um  $\bar{y}_{\text{DETECTION}}$ .

Bei  $\bar{y}_{\text{DETECTION}}$  liegt mit großer Sicherheit ein Analytsignal vor, das zur Bestimmung taugt – kleines Risiko  $\alpha$  aus Sicht der Signalverteilung an der Nachweisgrenze  $\bar{y}_{\text{CRIT}}$ . Dort ist andererseits – aus Sicht der Signalverteilung an der Bestimmungsgrenze  $\bar{y}_{\text{DETERM}}$  – das Risiko  $\beta$  gering, lediglich ein „Nachweisgrenzen-Signal“ ernst zu nehmen.

Historisch und auf zahlreichen Vereinfachungen begründet sind die sogenannten „ $k\sigma$ “-Kriterien, die in der Abbildung 10 erkennbar sind. Sie beruhen meist auf der Standardabweichung („ $\sigma$ “) von Leerwertmessungen: Bei  $k = 3$  definiert man die Nachweisgrenze, bei  $k = 6$  die Erfassungsgrenze und, als eine Möglichkeit, bei  $k = 10$  die Bestimmungsgrenze.<sup>44</sup>

Daraus ergibt sich die Abschätzung der *Erfassungsgrenze* als das Doppelte der Nachweisgrenze.

$$x_{\text{DETECTION}} \approx 2 x_{\text{DECISION}} \quad (67)$$

Dies entspricht zugleich dem Fall,  $\alpha = \beta$ ; für verschiedene Risiken,  $\alpha \neq \beta$ , ist die Erfassungsgrenze definiert als

$$x_{\text{DETECTION}} = x_{\text{DECISION}} + \frac{t(fg; q) s_{\text{RES}}}{a_1} \sqrt{\frac{1}{N_C} + \frac{1}{N_A} + \frac{\bar{x}_C^2}{SS_{XX}}} \quad (68)$$

wobei  $fg = N_L - 2$  und  $q = 1 - \beta$  ist.

Wie nach Abbildung 10 zu vermuten, kann für einen bestimmten Messwert  $y_n$  nur dann ein akzeptables Analytintervall angegeben werden, wenn die *Verfahrensstandardabweichung*  $s_{\text{RES}}/a_1$  nicht zu groß wird. Dieser Quotient ist eine kritische Größe für die praktische Anwendung eines Analysenverfahrens. Außerdem wird manchmal gefordert, dass die relative Standardabweichung eines Analytergebnisses einen vorgegebenen Maximalwert nicht übersteigt. Beide Forderungen führten zur Definition bzw. pragmatischen Festlegung von *Bestimmungsgrenzen*. Da solche Grenzwerte vom jeweiligen Messwert abhängen, können sie eigentlich nicht als sinnvolle Verfahrenskenngröße angesehen werden.<sup>45</sup>

44 G. L. LONG, J. D. WINEFORDNER: „Limit of detection; A closer look at the IUPAC definition“, Anal. Chem. 55(1983)712A-724A

45 Mittels der Verfahrensstandardabweichung schlägt DIN 32 645 (1994) auch Schnellschätzungen für die Nachweisgrenze und für die Bestimmungsgrenze vor.

Es muss darauf hingewiesen werden, dass die  $k$ -Werte die Wahrscheinlichkeit statistischer ( $\alpha$ ) und damit analytischer Fehlentscheidungen berücksichtigen. Somit kann durch Festlegen von  $k$  bzw.  $\alpha$  der Zweck des jeweiligen Spurenanalytikverfahrens flexibel berücksichtigt werden.

Die grundlegenden Überlegungen zu derartigen Definitionen von Verfahrensgrenzen gehen auf Arbeiten von H. KAISER (1973 und früher) und G. EHRLICH (1987) zurück. KAISER<sup>46</sup> versuchte u. a. auch, Verfahrenskenngrößen für die Mehrkomponentenanalytik zu definieren. (Man vergleiche dazu ggf. Abschnitt 5.3)

### 3.4 Die begriffliche Konfusion in der Literatur

Um die Vielfalt der Begriffe etwas zu illustrieren, sind in der folgenden Tabelle einige Varianten gegenübergestellt.

Referenz		$x_{\text{DECISION}}$	$x_{\text{DETECTION}}$	$x_{\text{DETERM}}$
Signalwert etwa bei	$y_{\text{CRIT}}$	$y_{\text{CRIT}} + 3\sigma$	$y_{\text{CRIT}} + 6\sigma$	$y_{\text{CRIT}} + 10\sigma$
<b>DIN 32645, 1994</b>	<b>krit. Wert der Messgröße</b>	<b>Nachweisgrenze</b>	<b>Erfassungsgrenze</b>	<b>Bestimmungsgrenze</b>
DIN 55350 <sup>47</sup>	Erkennungsgrenze	Erfassungsgrenze	Erfassungsvermögen	---
SCHWEDT <sup>48</sup>	---	Nachweisgrenze	Bestimmungsgrenze	---
MARR <sup>49</sup>	---	Erfassungsgrenze	Bestimmungsgrenze	---
MASSART <sup>50</sup>	critical value; decision limit; in terms of signals	decision limit (s. Anm. 1)	detection limit (s. Anm. 1)	determination; quantification limit (s. Anm. 1)
MILLER and MILLER, 1988	limit of decision	limit of decision	limit of detection (s. Anm. 2)	limit of quantitation determination
LUTHARDT <sup>51</sup>	Nachweisgrenze l.o. decision	Erfassungsgrenze l.o. detection	Bestimmungsgrenze l.o. determ.	---
CAULCUTT, 1983	criterion of detection	---	limit of detection	---
KAISER, 1973 (s. Anm. 4)	limit of detection	limit of detection (s. Anm. 5)	limit of guarantee / of purity	---
EHRLICH <sup>52</sup> 1970	---	Nachweisgrenze (s. Anm. 3)	Erfassungsgrenze	---

46 H. KAISER: „Zur Definition von Selektivität, Spezifität und Empfindlichkeit von Analysenverfahren“, Z. Anal. Chem. 260(1972) 252-260. Eine seiner ersten Arbeiten zu diesem Thema findet sich in Spectrochim. Acta 3(1947)40-67.

47 DIN 55 350, Teil 34 (Februar 1991): „Begriffe der Qualitätssicherung und Statistik; Erkennungsgrenze, Erfassungsgrenze und Erfassungsvermögen“, Beuth Verlag, Berlin, Köln

48 G. SCHWEDT: „Taschenatlas der Analytik“, Thieme Verlag, Stuttgart, 1992

49 I. L. MARR, M. S. CRESSER, L. J. OTTENDORFER: „Umweltanalytik“, Thieme Verlag, Stuttgart, 1988

50 D. L. MASSART, B. G. M. VANDEGINSTE, S. N. DEMING, Y. MICHOTTE, L. KAUFMAN: „Chemometrics: a textbook“, Elsevier, Amsterdam, 1988

51 M. LUTHARDT, E. THAN, H. HECKENDORFF: „Nachweis-, Erfassungs- und Bestimmungsgrenze analytischer Verfahren“, Fresenius J. Anal. Chem. 326(1987)331-339

52 G. EHRLICH, H.-G. DÖGE, A. DRESCHER, K. FRIEDRICH, R. GERBATSCH, H. MAI: „Spurenanalyse in hochschmelzenden Metallen“, Deutscher Verlag für Grundstoffindustrie, Leipzig, 1970

Anmerkungen zur Tabelle:

- (1) Alle Terme sind in der Signaldomäne definiert und nicht quantitativ (via Funktion) auf die Konzentrationsdomäne bezogen.
- (2) „In general terms, the limit of detection of an analyte may be described as that concentration which gives an instrument signal ( $y$ ) *significantly different* from the 'blank' or 'background' signal. It is immediately apparent that this description gives the analyst a good deal of freedom to decide the exact definition of the limit of detection, based on a suitable interpretation of the phrase 'significantly different'. There has in practice been very little agreement between professional and statutory bodies on this point, and the whole area is as controversial as other aspects of the statistical treatment of calibration analyses. ... It must be re-emphasized that this definition of a limit of detection is quite arbitrary, and is entirely open to an analyst to provide an alternative definition for a particular purpose.“ (MILLER and MILLER, 1988, p. 115f.)
- (3) „Die Angabe ... des ... Gehalts als Verfahrenskennwert ist nicht sinnvoll.“
- (4) und natürlich in seinen vorhergegangenen Schriften!
- (5) auf die Konzentration bezogen via  $c = f(\dots)$

## 4 Die Kalibrier-Techniken

---

### 4.1 Terminologie und Einführung

Die Konfusion der Begriffe in der „Kalibrierwelt“ ist erstaunlich. Wir werden versuchen, eine gewisse logische Namensgebung einzuhalten.

**Analyt** sei ein bekannter Stoff, von dem eine Analytgröße quantitativ bestimmt werden soll (*Beispiel*: Ammoniumion).

**Analytgröße** sei eine Eigenschaft des Analyten in einer bestimmten Probe. Während der Analyse werden Werte dieser Größe bestimmt. Wir verwenden dafür immer das Symbol  $X$  (*Beispiel*: meist die Konzentration).

**Probe** sei eine gewisse Menge eines Gases, Feststoffs oder einer Lösung, in der sich der Analyt befindet. Da diese Probe immer analysiert wird, ist das Synonym *Analysenprobe* akzeptabel.

Bei den Proben bzw. Analysenproben werden wir aber unterscheiden:

- (1) **Leerprobe** sei eine Probe, die den Analyten nicht enthält ( $x = x_{\text{LEER}} = 0$ ), aber dennoch dem gesamten Analysenprozess unterworfen wird.
- (2) **Kalibrierprobe** sei eine Probe, von der der Wert  $x = x_{\text{CAL}} = x_{\text{K}}$  der Analytgröße **bekannt** ist; unkorrekter Sprachgebrauch: „bekannte Probe“.
- (3) **Validierprobe** sei eine Probe, von der der Wert  $x = x_{\text{VAL}}$  der Analytgröße **bekannt** ist.
- (4) **Testprobe** sei eine Probe, von der der Wert  $x = x_{\text{TEST}} = x_{\text{A}}$  der Analytgröße **unbekannt** ist (unkorrekter Sprachgebrauch: „unbekannte Probe“).

**Standard-** oder **Referenzprobe** sei ein natürliches Gemisch oder eine Zubereitung, für die der Wert (die Werte) der Analytgröße (Analytgrößen) bekannt ist (sind). Standardproben sollten zertifiziert sein. Aus ihnen stellen wir Referenzproben her. Kalibrierproben sollten aus Standard- oder Referenzproben hergestellt sein.

In der Literatur sind folgende Synonyme bzw. Begriffe gebräuchlich:

Analyt	meint sowohl den Stoff als auch die Analytgröße.
Probe	meint meist die Testprobe.
Standard	meint häufig Kalibrierprobe oder aber auch den Analyten selbst (externer Standard) oder ganz und gar eine Fremdsubstanz (interner Standard).
Lernprobe	meint Kalibrierprobe.
Testprobe	meint häufig eine Validierprobe, mit der die Methode „getestet“ wird.
Blindprobe	meint Leerprobe.
Blindwert	meint Leerwert.
Eichen	meint den Vorgang des Kalibrierens.

(Der Begriff „Eichen“ sollte staatlichen Ämtern bzw. Behörden überlassen werden.)

Im englischen Schrifttum kann die Verwirrung noch krasser werden, zum Beispiel:  
sample heißt Probe im allgemeinen Sinn; auch Zufallsstichprobe.  
probe heißt „Sonde“ und nicht „Probe“.  
test sample heißt allgemein die Probe, die vermessen wird.

Bei den Techniken unterscheiden wir prinzipiell zwischen *externer Kalibrierung* und *interner Kalibrierung*.

Bei der üblichen externen Kalibrierung liegt eine Testlösung vor, in der ein Analyt vorliegt oder zumindest erwartet wird. Daraus prepariert man separate Kalibrierlösungen. Bei der internen Kalibrierung liegt immer zuerst die Testlösung mit Analyt vor. Zu ihr gibt man den Analyten in reiner Form (strenggenommen ist nur das eine „Analytzugabe“) oder in Form einer Kalibrierlösung, oder man entfernt den Analyt aus der Testlösung („Analytentfernung“).

Mit einer „internen Referenz“ ist meist kein Analyt gemeint.

In der Literatur wurden auch die Begriffe „sequentielle Kalibrierung“ („Separatkalibrierung“) und „simultane Kalibrierung“ verwendet. Sie beschreiben offenbar das Vorgehen bei der externen und der internen Kalibrierung.

Die in Abschnitt 2 ausführlich dargestellte Daten-Auswertemethode der *Regressionsanalyse* bezieht sich in der Regel auf die Technik der *externen Kalibrierung*. Sie ist allerdings universell anwendbar und kann auch bei den Techniken der Analytzugabe (siehe zum Beispiel Abschnitt 4.3.2), der internen Kalibrierung (siehe Abschnitt 4.7) und bei speziellen Techniken nützlich sein (siehe Abschnitt 4.8).

## 4.2 Externe Kalibration

### 4.2.1 ALLGEMEINES UND PRINZIP

#### *Synonyme:*

externer Standard; externe Kalibrierung; Kalibrierung

#### *Eignung:*

Diese Variante ist einzusetzen, wenn der Gesamtbereich modelliert werden soll, in dem voraussichtlich die Analytkonzentrationen liegen werden, und wenn sich der Aufwand für viele zukünftige Proben lohnt.

#### *Voraussetzungen:*

- (1) Ein geeigneter „Standard“ ist vorhanden. Im günstigsten Falle hat man ein Standard-Referenzmaterial, SRM, bzw. ein zertifiziertes Standard-Referenzmaterial, CSRM. Deren Matrix entspricht dem erwarteten Analyseproblem.
- (2) Eine lineare Beziehung zwischen Signal- und Analytgröße besteht wirklich.

#### *Vorteile:*

Diese Vorgehensweise ist intensiv untersucht und eine sehr gute statistische Absicherung ist möglich. Leider wird diese aber nicht immer genutzt.

### Nachteile:

Eine Nichtausführbarkeit ergibt sich dann, wenn keine geeignete Kalibrierprobe vorhanden ist bzw. nicht hergestellt werden kann. Das gilt besonders für die gleichzeitige Bestimmung mehrerer Analyte, denn für jede einzelne Komponente muss die Empfindlichkeit bestimmt werden.

### Prinzip der Vorgehensweise:

Im einfachsten Fall werden graphisch für eine Reihe von Kalibrierlösungen die gemessenen Analytsignale (Peakhöhen, -flächen, Stufenhöhen, Absorbanzen etc.) gegen die bekannten Werte der Analytgröße (Massen, Konzentrationen etc.) aufgetragen.

Die Punkte der Wertepaare werden linear verbunden bzw. es wird die Geradengleichung berechnet.

Nach diesem Kalibrierschritt wird im Analysenschritt das Analytsignal der Testprobe gemessen und im Kalibriergraph dazu der Wert der entsprechenden Analytgröße erfasst bzw. berechnet.

Das allgemeine Kalibriermodell hat die bekannte Form, siehe Gleichung (1), worin  $Y$  die Signalgröße und  $X$  die Analytgröße bedeuten.

### Anwendungsbeispiel Gaschromatographie (GC):

(A) Prinzipiell ist bei den Anwendungen in der GC die Art des Detektors zu berücksichtigen, damit im Kalibriermodell Signal- und Analytgröße „zusammenpassen“. So liefern FID, NPD, FPD und MSD massen- bzw. massenflussproportionale Signale ( $[X] = \text{m t}^{-1}$ ) und WLD (syn.: TCD, HWD), PID und IRD konzentrations- bzw. volumenstromproportionale Signale ( $[X] = \text{m v}^{-1}$ ).

Letztendlich kann jedoch ein Kalibriermodell unabhängig davon erstellt werden, solange die Betriebsbedingungen des Instruments konstant gehalten werden können und (für jede einzelne Analyt-Komponente) die Empfindlichkeit  $a_1$  bekannt und konstant ist. Natürlich ist zur Berechnung des Gehalts bzw. der Konzentration des Analyten die gefundene Analytmasse  $x_A$  auf die Masse bzw. das Volumen der Testprobe zu beziehen.

(B) Bei der *100%-Methode (Normierung auf 100%)* geht man davon aus, dass alle Komponenten im Chromatogramm erfasst wurden. Dann kann man den Anteil  $g$  (korrekt: den Volumenanteil) der  $c$ -ten Komponente direkt über folgende Gleichung berechnen:

$$g_c = \frac{y_c a_{1,c} 100}{\sum_{c=1}^c y_c a_{1,c}} \quad (69)$$

$g$	Volumenanteil in %
$y_c$	Peakfläche für Komponente $c$
$a_{1,c}$	(Detektor-)Empfindlichkeit für Komponente $c$

(C) Wenn eine reproduzierbare Probeninjektion gesichert ist, verfährt man häufig nach der inversen Variante (inverse Regression) und schätzt die Koeffizienten  $b_c$  für alle zu bestimmenden Komponenten C ( $c = 1, \dots, C$ ).<sup>53</sup>

#### 4.2.2 VEREINFACHUNG DURCH NUR EINMALIGE KALIBRIERPROBENMESSUNG

##### *Synonym:*

Einpunkt-Kalibrierung

Der Begriff ist insofern nicht richtig, als tatsächlich meist ein zweiter (Kalibrier-)Punkt nötig ist, zum Beispiel der Punkt (0; 0).

##### *Eignung:*

Diese Variante ist einsetzbar, wenn die lineare Kalibrierbeziehung nach Gleichung (1) gilt bzw. bekannt ist.

D. h., entweder ist eine Kalibrierbeziehung ohne  $a_0$  (keine Leerwerte) gegeben oder es gilt  $a_0 \neq 0$ , aber die Empfindlichkeit  $a_1$  ist bekannt.

Außerdem sollte die Analytkonzentration in der Kalibrierprobe sehr ähnlich der in der Testprobe sein, und es sollten sehr ähnliche Matrices vorliegen.

##### *Vorteil:*

Die Methode ist schnell.

##### *Nachteil:*

Eine evtl. Suche nach einer konzentrationsähnlichen Kalibrierprobe kann aufwändig sein.

##### *Vorgehensweise:*

- Analytsignal  $y_A$  der Testprobe registrieren
- Analytsignal  $y_K$  der Kalibrierprobe mit bekannter Konzentration  $x_K$  registrieren

##### *Auswertung:*

$$(1) \quad a_0 = 0$$

Wenn  $a_0 = 0$  ist, ergibt sich aus der direkten Relation (Streckenverhältnisse am Stufenwinkel) der Wert der Analytgröße in der Probe.

$$\frac{x_A}{y_A} = \frac{x_K}{y_K} \quad (70)$$

$$(2) \quad a_0 > 0$$

Wenn  $a_0 > 0$  ist, d. h. ein Leerwert  $y_{LEER}$  („Blindwert“) vorhanden und messbar ist, erhält man:

---

53 vgl. DIN 51 405: „Gaschromatographische Analyse, Allgemeine Arbeitsgrundlagen“, Beuth Verlag, Berlin, Köln, 1987. In dieser DIN wird sogar von „absoluter Kalibrierung“ gesprochen...

$$x_A = x_K \frac{y_A - y_{LEER}}{y_K - y_{LEER}} \quad (71)$$

(3)  $a_0 \neq 0$ ; jedoch nicht messbar

Wenn  $a_0 \neq 0$  ist, d. h. ein Leerwert  $y_{LEER}$  („Blindwert“) vorhanden, aber nicht messbar oder nicht gemessen ist, jedoch  $a_1$  bekannt ist, ergibt sich aus der Subtraktion der beiden Kalibrierungsdrücke für  $y_A$  und  $y_K$

$$x_A = x_K + \frac{y_A - y_K}{a_1} \quad (72)$$

*Anwendungsbeispiel Direktpotentiometrie:*

Hierbei verwendet man die Steigung der Kalibrierkurve explizit. Bei bekannter Steilheit  $S$  der Messkette ergibt sich mit  $\Delta U = U_A - U_K$

$$x_A = x_K + 10 \frac{\Delta U}{S} \quad (73)$$

## 4.3 Analyt-Zugabe zur Testprobe

### 4.3.1 ALLGEMEINES UND PRINZIP

*Synonyme:*

Analyt-Zugabe, Kalibrierzugabe, Aufstocken, Standardaddition, Inkrementtechnik, Zuzugabemethode, spiking, standard addition method, SAM, known addition

*Allgemeines:*

Die Analyt-Zugabemethode (AZM) wird dann empfohlen, wenn „Interferenzen durch die Matrix zu befürchten“ sind.

Genauer genommen sind das Interferenzen, die einen proportionalen systematischen Fehler bewirken würden. Das würde bedeuten: Der Parameter  $a_1$  des linearen Kalibriermodells gilt nicht im Analysenschritt.

Auch die AZM beruht auf der Annahme eines linearen Modells, setzt leerwertkorrigierte Messwerte und die Abwesenheit von Störungen (Interferenzen) anderer Analyte (Quersensitivität) voraus.

Eine solche Interferenz wäre im einfachsten Falle die Quelle eines additiven (konstanten) systematischen Fehlers, indem sie zu  $a_0$  beiträgt.

In solchen Fällen ist die verallgemeinerte Analyt-Zugabemethode (engl.: generalized standard addition method; GSAM) vorteilhaft, die üblicherweise dann angewendet wird,

wenn der störende Analyt explizit auch bestimmt werden soll. Dann ist eine Mehrkomponentenanalyse auszuführen, für die keine voll selektiven „Einzelkomponenten-Sensoren“ oder Messstellen zur Verfügung stehen.

Bei der inversen Voltammetrie sind derartige Störungen zum Beispiel durch im Quecksilber gebildete intermetallische Verbindungen vom Typ Zn-Cu berichtet worden.<sup>54</sup>

#### *Vorteil:*

Interessant ist die AZM vor allem dann, wenn mit einer einmaligen bzw. einzigartig strukturierten Probe zu rechnen ist, wofür Kalibrierproben naturgemäß schwer zu komponieren sind.

Der große Vorteil besteht hier darin, dass die Kalibrierung in der (Test-)Probe selbst stattfindet.

Vorteilhaft ist auch, dass Temperaturdifferenzen zwischen Kalibrierschritt und Anwendungsschritt ohne Einfluss sind bzw. nicht auftreten können.

Bei ISE-Messungen, wo die Kalibrierlösung immer im mindestens zehnfachen Überschuß vorliegt, sollten Matrixeffekte praktisch ausgeschlossen sein. (Ein TISAB-Puffer ist nicht nötig.)

#### *Nachteil:*

Bei Festkörperproben können durch Variation der Partikelzahl bzw. der Partikelgröße Inhomogenitäten auftreten.

Die ungefähre Analytkonzentration in der Testprobe muss bekannt sein.

Durch Zugabe (Verdünnen) ändert sich der Matrixeffekt. Deshalb ist wenigstens die Volumenzugabe möglichst gering zu halten.

#### *Prinzip:*

Es gibt zwei Varianten des Vorgehens, die man danach auswählen sollte, ob

- (1) das jeweilige Analysenverfahren den Analyten in der Analysenlösung (und die Lösung selbst) nicht merklich verbraucht, wie zum Beispiel in der Polarographie, oder ob
- (2) die Analysenlösung und damit der Analyt verbraucht wird, wie zum Beispiel bei der FAAS oder der AES mit ICAP.

Zur Illustration diene folgende Abbildung, in der A Analyt und Z Zugabe bedeutet, d. h. die X-Achse die Zugabemenge angibt.

---

54 R. W. GERLACH, B. R. KOWALSKI: „The Generalized Standard Addition Method: Intermetallic Interferences in Anodic Stripping Voltammetry“, Anal. Chim. Acta 134(1982)119-127

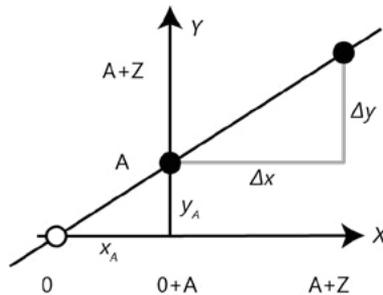


Abb. 11: Einfache Analytzugabe

### Vorgehensweise, allgemein:

(1)

- Registrieren des (leerwertkorrigierten) Signals von  $x_A$  in der Analysenlösung
- zur Analysenprobe (oder einem aliquoten Teil) nach und nach  $K$  gleiche Portionen  $v_Z$  einer Analytlösung der Konzentration  $x_Z$  geben
- jeweils das Signal von  $(x_A + k x_Z)$  registrieren ( $k = 1, 2, \dots, K$ )
- alle Signale gegen die Analytwerte auftragen
- Punkteschar linear ausgleichen
- die Gerade nach „negativen Zugabewerten“ verlängern
- beim Schnittpunkt mit der Abszisse den Wert von  $x_A$  ablesen

Diese Variante ist insbesondere dann einfach, wenn Volumenänderungen unberücksichtigt bleiben können. Bei der Variante (2) schließt man diese Volumenänderung direkt aus.

(2)

- $K$  Teilproben von der Analysenprobe entnehmen
- steigende Portionen einer Analytlösung zu den Teilproben  $k = 2, \dots, K$  geben
- Auffüllen aller Proben („Normieren“)
- Registrieren der Analytsignale aller Teilproben
- Auftragen aller Signalwerte gegen die Analytwerte
- Punkteschar linear ausgleichen
- beim Schnittpunkt der Kalibriergerade mit der Abszisse den Wert von  $x_A$  ablesen

### Vorgehensweise, quantitativ:

siehe dazu den folgenden Abschnitt

**Achtung:** Der prinzipielle Unterschied zwischen Mehrfach- und Einmal-Zugabe besteht darin, dass im ersten Fall der Analytanteil aus der Testprobe quasi als Leerwert angesehen wird.

Die Analytgröße im Regressionsmodell ist demnach nur aus der zugesetzten Analytmenge zu berechnen! Bei der Einmal-Zugabe wird dagegen mit der Gesamtmenge Analyt gearbeitet.

Genaugenommen dürfte man die Analyt-Zugabe-Methode im Wortsinne tatsächlich als interne Kalibrierung bezeichnen.

### 4.3.2 ... MIT MEHREREN ANALYT-ZUGABEN

#### Anwendungsbeispiel Polarographie:

- im Messgefäß befindet sich eine Grundlösung (Leitelektrolyt, Puffer, ...) mit  $v_G$
- dazu exaktes Volumen  $v_A$  der Testprobe der Konzentration  $x_A$  geben
- Gesamtvolumen ist nun  $v_0$ ; Konzentration des Analyten ist  $x'_A$
- Signal des Analyten messen ( $y_A = y_0$ )
- exaktes Volumen  $v_Z$  Analytlösung (des „Standards“) der Konzentration  $x_Z$  zugeben; Analytgesamtkonzentration ist jetzt  $x_1$
- Messen des Signals  $y_1$
- weitere Zugabe des exakten Volumens  $v_Z$  usw. bis  $K$  erreicht ist; jeweils  $y_k$  messen
- alle Wertepaare graphisch auftragen

Die Werte der (Stoffmengen-)Konzentration  $X_Z$  des Analyten an den Stellen  $k = 1, \dots, K$  berechnen sich einfach nach

$$x_k = \frac{k v_Z x_Z}{v_0 + k v_Z} \quad (74)$$

$v_Z$	Volumen der Lösung des Zugabeanalyten
$v_0$	Volumen im Messgefäß vor Zugabe ( $v_0 = v_G + v_A$ )
$x_Z$	Konzentration der Lösung des Zugabeanalyten

- Auswertvariante: graphisch
  - Gerade zeichnen
  - $x'_A$  direkt bei  $Y = f(X_Z) = 0$  ablesen
- Auswertvariante: Regression
  - $a_0$  und  $a_1$  für die Gleichung  $Y = f(X_Z)$  berechnen, beachten, dass prinzipiell  $a_0 \neq y_A$  ist
  - $x'_A$  berechnen nach

$$x'_A = \frac{a_0}{a_1} \quad (75)$$

- Die Konzentration des Analyten in der Teilmenge  $v_A$  der Testprobe berechnen nach

$$x_A = x'_A \frac{v_0}{v_A} \quad (76)$$

Der Fehler von  $x_A$  berechnet sich nach SUNDBERG<sup>55</sup> näherungsweise aus

$$s_A^2 = \left( \frac{a_0}{a_1} \right)^2 \left( \frac{1}{K+1} + \frac{(x_A + \bar{x}_K)^2}{SS_{XX}} \right) \quad (77)$$

wobei sich  $SS_{XX}$  von  $k = 0$  bis  $K$  erstreckt. TIMM et al.<sup>56</sup> geben das statistisch exakte Konfidenzintervall für den Fall an, dass für jede Additionsstufe eine gleiche Anzahl von Parallelmessungen gemacht wurde.

### 4.3.3 ... MIT EINMALIGER ANALYT-ZUGABE

Dieser Spezialfall läßt sich aus dem Fall der mehrmaligen Zugabe schnell ableiten, wenn man, statt die Koeffizienten  $a_0$  und  $a_1$  regressionsanalytisch zu schätzen, von folgender Relation ausgeht, also streng linearen Verlauf der Kalibrierkurve unterhalb der Konzentration  $x_1$  annimmt, die Kurve direkt durch  $(0, y_A)$  und  $(x_1, y_1)$  zeichnet ( $a_0 = y_A$  !):

$$\frac{a_0}{a_1} \approx \frac{y_A}{\Delta y / x_1} = \frac{y_A}{\Delta y} x_Z \frac{v_Z}{v_0 + v_Z} \quad (78)$$

$\Delta y = y_1 - y_A$  ist die Differenz der Signalwerte nach und vor der Zugabe.

Damit hat man sofort  $x'_A$  und kann wie oben  $x_A$  berechnen.

Außerdem kann bei vernachlässigbarer Volumenzunahme die Berechnung von  $x_A$  noch etwas vereinfacht werden:

$$x_A \approx \frac{y_A}{\Delta y} x_Z \frac{v_Z}{v_A} \quad (79)$$

Verwendet man für die Probenzugabe und die Analytzugabe zur Grundlösung das gleiche geringe Volumen, läßt sich sogar noch eine weitere Vereinfachung machen, mit der in der Praxis häufig gerechnet wird: Dann geht man nämlich von der simplen Relation aus, dass

$$\frac{x_A}{y_A} \approx \frac{x_Z}{\Delta y} \quad (80)$$

55 R. SUNDBERG: „Interplay between Chemistry and Statistics, with Special Reference to Calibration and the Generalized Standard Addition Method“, Chemom. Intell. Lab. Sys. 4(1988)299-305

56 J. TIMM, H. DIEHL, D. HARBACH: „Voraussetzungen und Grenzen zur Anwendung der Additionsmethode bei der Atomabsorptions-Spektrometrie“, Fresenius Z. Anal. Chem. 301(1980)199-202

Mit anderen Worten, man könnte eigentlich zwei Analysen ausführen, eine mit dem Probevolumen und eine mit dem Analytzugabevolumen, und wäre dann wieder bei der Relation, die exakt bei der externen (üblichen) Kalibrierung im Falle  $a_0 = 0$  gilt! Allerdings zeigt sich hier auch klar, dass man sehr darauf achten muß, dass beide Konzentrationen eng beieinander liegen müssen.

#### Anwendungsbeispiel Polarographie:

Letztgenannte Formel findet in der Polarographie Anwendung. Man beachte aber die notwendigen Vereinfachungen, die zu dieser Formel führten. Man vergleiche außerdem mit dem Ergebnis aus Abschnitt 4.2.2.

#### Anwendungsbeispiel Direktpotentiometrie:

- Elektrodensteilheit  $S$  bestimmen
- Testprobenvolumen  $v_A$  mit  $x_A$  vorlegen; ggf.  $v_G$  TISAB o. ä. zugeben, so dass  $x'_A$  bezogen auf Gesamtvolumen  $x_0$  vorliegt; anderenfalls ist  $x_0$  gleich  $x_A$
- Potential  $U_A$  messen
- Kalibrierprobenvolumen  $v_Z$  mit  $x_Z$  zusetzen; Gesamtkonzentration des Analyten wird  $x_1$  bezogen auf  $v_0 + v_Z$
- Potential  $U_1$  messen; Differenz berechnen:  $\Delta U = U_1 - U_A$
- Analytkonzentration in der Testprobe berechnen nach der Überlegung:

$$\begin{aligned} U_1 &= U_0 + S \lg(x_1) \\ U_A &= U_0 + S \lg(x'_A) \\ \Delta U &= S \lg\left(\frac{x_1}{x'_A}\right) \end{aligned} \quad (81)$$

woraus sich mit

$$x_0 = x'_A = x_A \frac{v_A}{v_0} \quad x_1 = \frac{v_A x_A + v_Z x_Z}{v_0 + v_Z} \quad (82)$$

und den Volumenverhältnissen  $Q_0$  bzw.  $Q_V$  und dem Potentialterm  $T$

$$Q_0 = \frac{v_0}{v_0 + v_Z} \quad Q_V = \frac{v_Z}{v_A} \quad T = 10^{\frac{\Delta U}{S}} \quad (83)$$

als Ergebnis für die Messprobe  $x'_A$  ergibt:

$$x'_A = x_Z \frac{v_Z}{v_A} \frac{Q_0}{T - Q_0} \quad (84)$$

Daraus kann  $x_A$  ermittelt werden, wenn man  $v_0 = v_A$  annimmt:

$$x_A = x_Z \frac{v_Z}{v_A + v_Z} \frac{1}{T - \frac{v_A}{v_A + v_Z}} \quad (85)$$

Unter der gleichen Voraussetzung erhält man in anderer Schreibweise:<sup>57</sup>

$$x_A = x_Z \frac{Q_V}{(1 + Q_V) T - 1} \quad (86)$$

Alle in der Literatur zu findenden „bequemeren“ Formeln beruhen auf bestimmten festen Verhältnissen der Volumina bzw. der Annahme idealer Steilheit für ein bestimmtes Mession.

#### 4.3.4 ... MIT DOPPELTER ANALYT-ZUGABE

Diese Methode wird manchmal beim Arbeiten mit ionensensitiven Elektroden (ISE) empfohlen, insbesondere dann, wenn in Bereichen gearbeitet werden soll, wo ein nichtlineares Ansprechverhalten der ISE vorliegt.

Bei Messungen mit der Ca-ISE muss insbesondere bei Proben mit hohen Ionenstärkewerten (wesentlich größer als  $0,01 \text{ mol l}^{-1}$ ) eine größere Menge TISAB zugesetzt werden. Bei den dann entstehenden hohen Salzgehalten sollte nach der Analytzusatzmethode gearbeitet werden.

Wahrscheinlich ist sie genau dann die Methode der Wahl, wenn  $a_1$  nicht konstant ist.

### 4.4 Zugabe der Testprobe zu einer Kalibrierprobe

*Synonyme:*

Testproben-Zugabe, Probenaddition, Inkrementtechnik, analyte addition

*Eignung:*

Diese Vorgehensweise kann man einsetzen für Testproben mit hohen Konzentrationen, bei denen die übliche Analyt-Zugabe nur kleine Signaländerungen ergäbe.

*Vorteil:*

Auch Proben mit sehr kleinen Volumina lassen sich so analysieren.

---

57 Firmenschrift, ORION: „Ion-Selective Electrode Catalog and Guide to Ion Analysis“, 1992

### Anwendungsbeispiel Potentiometrie mit ISE.<sup>58</sup>

- Elektrodensteilheit bestimmen
- Kalibrierprobe mit Volumen  $v_Z$  entsprechend erwarteter Analytgröße in der Testprobe vorlegen
- Potential messen
- kleines Volumen  $v_A$  der Testprobe zusetzen
- Potential messen; Differenz bestimmen
- Analytkonzentration  $x_A$  in der Probe berechnen:

$$x_A = x_Z \frac{v_Z + v_A}{v_A} \left( T - \frac{v_Z}{v_A + v_Z} \right) \quad (87)$$

Unter der gleichen Voraussetzung wie oben erhält man in anderer Schreibweise:<sup>59</sup>

$$x_A = x_Z \left( (1 + Q_V) T - Q_V \right) \quad (88)$$

## 4.5 Analyt-Entfernung aus einer Testprobe

### Synonyme:

Analyt-Subtraktion, Standardsubtraktion, Inkrementtechnik, known subtraction

### Vorteil:

Dieser besteht in Fällen, wo Kalibrierproben instabil sind bzw. darin keine Konzentrationen bestimmt werden können.

### Anwendungsbeispiel Direktpotentiometrie:

(z. B. Cyanidbestimmung; Reagens: Hypochlorid, das Cyanid oxidiert)

- Elektrodensteilheit bestimmen
- Testprobe mit Volumen  $v_A$  vorlegen
- Potential messen
- Reagensvolumen  $v_Z$  zusetzen; Reagens reagiert mit Analyt und soll dessen Konzentration lediglich *verringern*
- Potential messen; Differenz bestimmen
- Analytkonzentration in der Testprobe berechnen

58 U. GRÜHNKE, P. HARTMANN: „Ionensensitive Festkörpermembran-Elektroden -- Eigenschaften und Anwendungsmöglichkeiten“, Hermsdorfer Techn. Mitt. 18(1978)1619-1654, Heft 51; Firmenschrift, METTLER-TOLEDO, Prozessanalytik GmbH: „Anleitung zur ionenselektiven Messung“, 1996

59 Firmenschrift, ORION: „Ion-Selective Electrode Catalog and Guide to Ion Analysis“, 1992

Unter der gleichen Voraussetzung wie oben erhält man in anderer Schreibweise:<sup>60</sup>

$$x_A = x_Z \frac{Q_V}{1 - (1 + Q_V) T} \quad (89)$$

## 4.6 Analyt-Entfernung aus einer Kalibrierprobe

*Synonyme:*

Analyt-Subtraktion, analyte subtraction, Probensubtraktion, Inkrementtechnik  
(Analyt-Entfernung durch Zugabe einer Testprobe)

*Vorteil:*

Das Verfahren entspricht einer indirekten Bestimmungsweise. Es können Analyte bestimmt werden, für die kein Sensor existiert.

*Anwendungsbeispiel Direktpotentiometrie:*

(z. B. Sulfatbestimmung; Reagens: Bariumchlorid, das Sulfat fällt)

- Elektrodensteilheit bestimmen
- Kalibrierprobe mit Mession (Reagens) vorlegen
- Potential messen
- kleines Volumen Testprobe zusetzen; Analyt der Testprobe reagiert mit Analyt der Messprobe und verringert dessen Konzentration
- Potential messen; Differenz bestimmen
- Analytkonzentration in der Testprobe berechnen

Unter der gleichen Voraussetzung wie oben erhält man in anderer Schreibweise:<sup>61</sup>

$$x_A = x_Z \left( Q_V - (1 + Q_V) T \right) \quad (90)$$

---

60 Firmenschrift, ORION: „Ion-Selective Electrode Catalog and Guide to Ion Analysis“, 1992

61 Ebd.

## 4.7 Interne Kalibrierung

### 4.7.1 PRINZIP UND VORGEHEN

#### Synonyme:

Interner Standard, interne Referenz, internal standard, internal reference

#### Eignung:

Die Methode wird vor allem verwendet, wenn keine sichere Reproduzierbarkeit der Dosierung möglich ist. Dies kann bei chromatographischen Verfahren auftreten.

#### Vorteil:

Die interne Kalibrierung ist universell einsetzbar und zur Spurenanalytik geeignet, da Fehler bei der Probeninjektion minimal gehalten werden können.

#### Nachteile:

Es gibt keinen universell einsetzbaren Standards. An die Kalibriersubstanz und den internen Standard werden sehr hohe Reinheitsanforderungen gestellt.

#### Vorgehen:

Hierbei fügt man sowohl den Kalibrierproben als auch der Testprobe eine *weitere Komponente* hinzu, den „internen Standard“, d. h. eine Komponente, die vorher nicht in den Proben enthalten war.

In der Chromatographie sollte diese Referenzsubstanz basislinienaufgelöst nahe dem Analyten eluierbar sein.

Zur Ermittlung der Kalibrierkurve wird das Signalverhältnis Analyt/Referenz gegen das Massenverhältnis Analyt/Referenz aufgetragen (graphische Variante), bzw. werden die Koeffizienten im allgemeinen Kalibriermodell berechnet:

$$\frac{Y}{Y_{REF}} = b_0 + b_1 \frac{X}{x_{REF}} \quad (91)$$

(Man beachte hier die Bedeutung von Groß- und Kleinschreibung.)

Alternativ ist auch folgende Schreibweise mit Flächenkorrekturfaktoren  $f$  möglich:

$$\frac{Y_{REF}}{Y} = f_0 + f_1 \frac{x_{REF}}{X} \quad (92)$$

Der Wert  $x_{REF}$  der Referenzsubstanzgröße (Masse) muss konstant sein.

Im Analysenschritt wird das Signalverhältnis von Analyt/Referenz bestimmt und aus dem abgelesenen Massenverhältnis Analyt/Referenz mit der bekannten Masse der Referenz  $x_{REF}$  die Analytmasse berechnet.

$$\hat{x}_A = \frac{x_{REF}}{b_1} \left( \frac{y_A}{y_{REF}} - b_0 \right) \quad (93)$$

#### Anwendungsbeispiel Chromatographie:

Im Allgemeinen geht man von basislinienkorrigierten Messwerten aus, d. h. bei konstanter „Nulllinie“ tritt kein Leerwert in obigen Gleichungen auf.

Mit der Ermittlung des Anstiegs  $b_1$  ist zugleich der *response factor* bestimmt.

Der Kehrwert dieses Faktors wird auch manchmal als Flächenkorrekturfaktor  $f_1$  bezeichnet.

#### 4.7.2 VEREINFACHUNGEN

Wenn der Anstieg  $b_1$  der Kalibrierkurve konstant ist bzw. als konstant angenommen werden kann, ist statt der oben beschriebenen *Mehrpunktkalibrierung* ein *Zweischrittverfahren* möglich:

(1)

- eine Kalibrierprobe mit Analyt  $x_{K,A}$  und mit interner Referenz  $x_{K,REF}$  herstellen
- chromatographieren, Signalwerte  $y_{K,A}$  und  $y_{K,REF}$  ermitteln
- ganz im Rahmen eines der obigen Modelle berechnen, zum Beispiel:

$$\frac{y_{K,REF}}{y_{K,A}} / \frac{x_{K,REF}}{x_{K,A}} = f_1 \quad (94)$$

(2)

- die (eine) Testprobe, die  $x_A$  enthält, mit interner Referenz  $x_{REF}$  versetzen
- chromatographieren, Signalwerte  $y_A$  und  $y_{REF}$  ermitteln
- berechnen:

$$x_A = f_1 x_{REF} \frac{y_A}{y_{REF}} \quad (95)$$

### Anwendungsbeispiel HPLC:

Nach dem geschilderten Muster verfährt man in der HPLC, wobei  $b_1$  als *Responsefaktor* (engl.: response factor) bzw.  $f_1$  als *Korrekturfaktor* bezeichnet werden sollte und natürlich für jede einzelne Komponente zu berechnen ist.

## 4.8 Weitere „Spezialfälle“

### 4.8.1 MEHRFACHPROBENAHE AUS DEM DAMPFRAUM (GC)

#### Synonyme:

Multiple Headspace-Extraktion (MHE), Mehrfachextraktion aus dem Dampfraum

#### Eignung:

Die Kalibrierung bei festen Proben ist äußerst problematisch und praktisch fast unmöglich.

Die MHE ist eine mögliche Technik, die Festproben-Kalibrierung zu umgehen. Sie ist jedoch selbst keine Kalibriertechnik, sondern ihr Ziel ist es, die für die Kalibrierung nötigen Signalwerte (Peakflächen) in besonderer Weise zu beschaffen.

Sie beruht bei Vorliegen eines Gleichgewichts zwischen dem Analyten in der festen Matrix (oder der „modifizierten“ Matrix, siehe unten) auf der exponentiellen Abnahme des Signals in jedem Analysenschritt  $s$  nach

$$y_s = y_1 \exp(- (s - 1) k^*) \quad (96)$$

Bei logarithmischer Darstellung  $\ln y_s$  gegen  $s$  ist der Anstieg zwischen den beiden ersten Schritten

$$k^* = \ln \frac{y_1}{y_2} \quad (97)$$

Daraus ergibt sich mittels der Grenzformel für eine geometrische Reihe das Signal nach vollständiger Extraktion des Analyten zu

$$y_{sum} = \frac{y_1}{1 - \exp(-k^*)} \quad (98)$$

#### Allgemeines Vorgehen:

(HS = Headspace, Dampfraum)

- Vorlegen der festen oder flüssigen Testprobe im HS-Gläschen

- Probe ggf. mit einem *modifier* versetzen, der ein geeignetes Phasenverhältnis des Analyten im HS-Gefäß sichern soll
- Verschließen und Temperieren des HS-Gefäßes: Äquilibrieren der Probe
- mehrfache Entnahme einer Portion aus dem Dampfraum mit jeweils anschließender Gleichgewichtseinstellung
- nach jeder Entnahme erfolgt eine Chromatogramm-Aufnahme und die Berechnung der schrittweise erhaltenen Peakflächen  $y_{c,s}$  (Schrittnummer:  $s = 1, 2, \dots$ )
- für jeden Analyt Auftragen von  $\ln(y_{c,s})$  gegen Schrittnummer  $s$
- Berechnen der Flächensumme  $y_{c,sum}$ 
  - nach Zweischrittvariante aus:

$$y_{c,sum} = \frac{y_1^2}{y_1 - y_2} \quad (99)$$

(„Graphische Lösung“ durch Einsetzen des Ausdrucks für  $k^*$  in die Formel für  $y_{sum}$ )

- oder nach Regressionsrechnung zur Ermittlung von

$$\hat{y}_1 \quad \text{und} \quad \hat{k}^* \quad (100)$$

und Einsetzen der beiden Werte in die Formel für  $y_{sum}$

Es sind folgende Kalibriervarianten denkbar, die jeweils zu  $y_{c,sum,K}$  führen:

(1) externe Kalibrierung mit einer Kalibrierprobe:

- eine Leerprobe plus Mischung der reinen Analyte
- die Testprobe
- Berechnung der Masse  $m_{c,A}$  des Analyten  $c$  in der Testprobe A aus dem Dreisatz auf Basis von:

$$\begin{aligned} y_{c,sum,K} &= m_{c,K} \\ y_{c,sum,A} &= m_{c,A} \end{aligned} \quad (101)$$

(d. h. hier sind zwei Analysen nötig und zwei HS-Gläschen einzusetzen)

(2) externe Kalibrierung mit mehreren Kalibrierproben und Ablesen bzw. Berechnen der Analytmasse wie üblich

#### 4.8.2 ELISA-KALIBRIERUNG

Bei enzymatischen Analysenmethoden, wie zum Beispiel dem ELISA (engl.: enzyme linked immuno sorbent assay), erhält man meist *sigmoidale* Abhängigkeiten des Signals vom (natürlichen) Logarithmus der Konzentration des Analyten (log Dosis-Response-Kurve). Mit der zwischen 0 und 1 normierten Signalgröße

$$Y_{norm} = \frac{Y}{Y_0} \quad (102)$$

wird die nichtlineare Signal-Konzentrationsbeziehung in die lineare logit-log-Funktion

$$\text{logit} (Y_{norm}) = a_0 + a_1 \ln X \quad (103)$$

umgewandelt.

Aus der Definition

$$\text{logit} (Y_{norm}) = \ln \frac{Y_{norm}}{1 - Y_{norm}} \quad (104)$$

ergibt sich

$$\ln \frac{Y}{Y_0 - Y} = a_0 + a_1 \ln X \quad (105)$$

$Y$	Signal (des gebundenen Tracers) in Anwesenheit des Analyten
$Y_0$	Signal bei Abwesenheit des Analyten

Mit dieser Beziehung kann wie üblich kalibriert werden.

#### 4.9 Vorläufige Orientierung: Welche Technik wann anwenden?

Zumindest in der GC kann folgendes Frageschema zur richtigen Wahl der Kalibriertechnik führen:

- (1) Reiner Analyt verfügbar?
  - JA, dann weiter bei (2).
  - NEIN, dann aufgeben. (Es sei denn, man findet einen internen Standard, der eine vergleichbare Empfindlichkeit wie der Analyt aufweist.)

(2) Quasi leere Matrix herstellbar oder verfügbar?

- JA, dann externe Kalibrierung und weiter bei (3).
- NEIN, dann Analytzugabe-Technik und weiter bei (3).

(3) Ist mit Dosierungsproblemen (Reproduzierbarkeitsproblemen) zu rechnen?

- JA, dann Technik mit interner Referenz koppeln.
- NEIN, dann Technik wie üblich anwenden.

## B Mehrkomponentenanalytik

### 5 Zweikomponenten-Kalibration

---

#### 5.1 Modellgleichung, Signalmodellierung

Im Einkomponenten-Fall sind wir von dem allgemeinen Modell nach Gleichung (1) ausgegangen, dessen Symbolik wir für diesem Teil geringfügig ändern:

$$Y = b_0 + b_1 X_1 \quad (106)$$

Die Änderung der Symbole ist zwar nicht zwingend nötig, wir wollen damit nur den Übergang zur multivariaten Kalibrierung einleiten.

$X_1$  ist wieder die veränderliche „Analytgröße“, also eine Analyteigenschaft wie zum Beispiel die Konzentration des Analyten. Den Beitrag  $b_0$  zur Signalgröße  $Y$ , in der Fotometrie etwa als „Leerwert“ bezeichnet, könnten wir auch als konstanten Beitrag einer (unbekannten) Störkomponente verstehen. Sofern dieser Beitrag konstant bleibt, gibt es keine Probleme.

Wenn er jedoch von der Konzentration der Störkomponente veränderlich abhängt, muss diese bekannt sein und als signalgebende Größe  $X_2$  explizit berücksichtigt werden.

Die Störkomponente avanciert damit zum zweiten Analyten:

$$Y = b_1 X_1 + b_2 X_2 \quad (107)$$

Um beide Koeffizienten dieses univariaten Modells berechnen zu können, sind mindestens zwei Bestimmungsgleichungen nötig. Praktisch hat man dazu zwei Möglichkeiten.

(1) Messung nur einer Signalgröße

Das Vorgehen entspricht dann etwa dem in der statistischen Versuchsplanung.

Dabei variiert man im Kalibrierschritt die Anteile der Einflussfaktoren, nämlich der beiden Größen  $X_1$  und  $X_2$  in den Kalibrierproben, ermittelt die Werte der Signalgröße, zum Beispiel die Absorbanz bei *einer* Wellenlänge  $\lambda_{MAX}$ , und findet mittels Regressionsrechnung die beiden Koeffizienten.

Meist kann die Koeffizientenberechnung sogar per Hand erfolgen.

Es genügen zwei Bestimmungsgleichungen:

$$\begin{aligned} y_1(\lambda_{MAX}) &= b_1(\lambda_{MAX}) x_{11} + b_2(\lambda_{MAX}) x_{12} \\ y_2(\lambda_{MAX}) &= b_1(\lambda_{MAX}) x_{21} + b_2(\lambda_{MAX}) x_{22} \end{aligned} \quad (108)$$

Zwei Kalibrierproben mit jeweils unterschiedlichen Konzentrationen von  $X_1$  und  $X_2$  wären also schon ausreichend. Statistisch vorteilhafter ist natürlich die Verwendung von  $N$  Kalibrierproben, deren Konzentrationen zum Beispiel nach einem *orthogonalen Versuchsplan* variieren.

Vielleicht muss man außerdem Leerwerte berücksichtigen, die analytisch wie immer bedingt sein können. In unserem Datenbeispiel werden wir sie selbstverständlich auch berechnen.

❶ Datenbeispiel siehe Abschnitt 5.7

(2) Messung mehrerer Signalgrößen

In Hinblick auf die häufigste praktische Anwendungsvariante, zum Beispiel in der photometrischen Mehrkomponenten-Analyse, besprechen wir diesen Fall bereits hier ausführlicher.

Im Prinzip braucht man noch einen zweiten „Messkanal“, einen zweiten Sensor. Der erste Sensor sollte hauptsächlich den Analyten  $X_1$  erfassen, der zweite hauptsächlich  $X_2$ . In der photometrischen Zweikomponenten-Analyse würde man dazu einfach die gleiche Messgröße Absorbanz bei zwei verschiedenen Wellenlängen messen.

Der Vorteil dieses Weges ist, dass man statt des Analyt-Verhältnisses (siehe Versuchsplanung) einfach die Messstelle variiert. Der praktische Aufwand ist dadurch deutlich geringer, auch wenn man im einfachsten Fall zwei Kalibrierproben mit unterschiedlichen Analytwerten benötigt.

Diese Situation wird durch folgende Gleichungen beschrieben:

$$\begin{aligned} Y(\lambda_1) &= b_1(\lambda_1) X_1 + b_2(\lambda_1) X_2 \\ Y(\lambda_2) &= b_1(\lambda_2) X_1 + b_2(\lambda_2) X_2 \end{aligned} \tag{109}$$

Wir haben damit praktisch zwei unabhängige univariate Modellgleichungen zu einem bivariaten Gleichungssystem für zwei Analyt-Faktoren vereinigt.

Dieser Schritt ist lediglich deshalb erforderlich, weil wir im Anwendungsschritt zwei Bestimmungsgleichungen zur Bestimmung zweier Analyte in einer Probe benötigen.

Dieser Anwendungsfall tritt sowohl bei Validier- als auch bei Analysenproben ein.

Formal würden sich in den folgenden Definitionen lediglich bestimmte Matrix- bzw. Vektor-Dimensionen ändern. Um die weiteren Darstellungen aber nicht zu komplizieren, werden wir diesen Aspekt nicht einbeziehen. Man kann sich auch einfach vorstellen, dass man es mit Leerwert-korrigierten Signalen zu tun hat.

Schreiben wir das Gleichungssystem (ohne die Leerwerte) nun als Wertgleichungssystem und bedienen uns der Symbolik der Matrix-Algebra, so können wir es mit  $\mathbf{y}$ ,  $\mathbf{B}$  und  $\mathbf{x}$  als (2,1)-, (2,2)-, (2,1)-dimensionale Vektoren bzw. Matrizen

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (110)$$

kurz schreiben als:

$$\mathbf{y} = \mathbf{B} \mathbf{x} \quad (111)$$

In dieser Schreibweise kann man bequem zu einem Mehrkomponentensystem übergehen, indem man mit  $(P,1)$ -,  $(P,C)$ -,  $(C,1)$ -dimensionalen Vektoren bzw. Matrizen arbeitet, wobei  $P$  die Anzahl der Signal- oder Messgrößen (Sensoren, Messkanäle,...) oder der „Messstellen“ einer Signalgröße ist und  $C$  die Anzahl der Komponenten angibt.

Wenn  $P = C$  ist, spricht man von exakt bestimmten Gleichungssystemen, für  $P > C$  liegt ein überbestimmtes Gleichungssystem vor. Bei Analysenmethoden, die das komplette Spektrum eines Analyten registrieren, hat man es also in der Regel mit überbestimmten Systemen zu tun.

Für den Fall  $P = C$  hat KAISER (1972) den Begriff der partiellen Empfindlichkeiten eingeführt.

Sie ergeben sich als

$$b_{pc} = \frac{\partial Y_p}{\partial x_c} \quad (112)$$

Damit enthält die Kalibriermatrix  $\mathbf{B}$  die partiellen Empfindlichkeiten  $b_{pc}$  aller  $P$  Messstellen bezüglich jeder der  $C$  Komponenten, und das Ermitteln von  $\mathbf{B}$  nennt man Kalibrieren.

Um eine sprachliche Irritation durch die Verwendung des Begriffs Matrix für das stoffliche Medium des oder der Analyten zu vermeiden, wird fortan der etwas sperrige Begriff „Empfindlichkeitenmatrix“ verwendet.

## 5.2 Modellparameter, Empfindlichkeitenmatrix

Manche Autoren unterscheiden direkte und indirekte Kalibrierung. Sie verstehen unter Kalibrieren nämlich sowohl den Kalibrier- als auch den Analyseschritt. Aus methodischen Gründen betrachten wir diese beiden Schritte getrennt.

Bei der direkten Kalibrierung ist die Empfindlichkeitenmatrix vor der Analyse bekannt, zum Beispiel in Form der Spektren der reinen Komponenten.

- ❶ Man vergleiche das Datenbeispiel zur direkten Kalibration in Abschnitt 6.4.1

Bei der indirekten Kalibrierung, für das übliche Vorgehen also, wird im Kalibrierschritt zunächst die Empfindlichkeitenmatrix  $\mathbf{B}$  ermittelt. Dazu sind in der Regel  $N$  Kalibrierproben (Versuche) nötig, in denen die  $P$  Komponenten (Analyte) in  $C$  unterschiedlichen Konzentrationen (Anteilen) vorliegen.

Im Allgemeinen wählt man die  $N$  Versuche zufällig aus und kann dann auch von zufälliger Kalibrierung sprechen.

Günstiger ist jedoch immer eine systematische Variation der Analytwerte aller Komponenten (Konzentrationen), im Optimalfall nach einem statistischen Versuchsplan. Dann liegt eine *kontrollierte Kalibrierung* vor.

Bezüglich der Vorhersage der Konzentrationen implizieren diese beiden Wege unterschiedliche statistische Eigenschaften der ermittelten Konzentrationen (NAES, 1985).

## 5.3 Bewertung und Auswahl eines Kalibrationsmodells

### 5.3.1 ... NACH DER EMPFINDLICHKEIT

Im Einkomponenten-Fall war die Messung nur einer Signalgröße ausreichend, zum Beispiel bei nur einer Wellenlänge. Die Empfindlichkeit war mit der partiellen Empfindlichkeit für den jeweiligen Analyten identisch.

Für den Zweikomponenten-Fall, aber auch im Mehrkomponenten-Fall, würde man im Idealfall eine quadratische ( $C,C$ )-Empfindlichkeitenmatrix  $\mathbf{B}$  haben, deren Hauptdiagonalelemente absolut dominant sind.

In der Praxis heißt das zum Beispiel, dass man aus den Spektren der reinen Analyte jeweils diejenige Messstelle  $\lambda_{\text{MAX}}$  auswählen sollte, wo die Absorption maximal ist.

Diese Variante ist natürlich nur dann anwendbar, wenn die Empfindlichkeitsmaxima der Analyte, d. h. deren Messstellen, nicht zusammenfallen.

Für Mehrkomponenten-Systeme, wo zum Beispiel Messungen bei mehreren Messstellen (zum Beispiel Wellenlängen) erforderlich sind, hat KAISER, wieder für den Fall exakt bestimmter Kalibriersysteme,  $P = C$ , als analoge Kenngröße für die Empfindlichkeit ( $sen_{CC}$ ) die Determinante der Empfindlichkeitenmatrix vorgeschlagen:

$$sen_{CC} = |\det(\mathbf{B})| \quad (113)$$

$\mathbf{B}$  (C,C)-Matrix der partiellen Empfindlichkeiten

Für überbestimmte Systeme,  $P > C$ , kann man folgende Empfindlichkeitszahl<sup>62</sup> ( $sen_{PC}$ ) berechnen:

$$sen_{PC} = \sqrt{\det(\mathbf{B}^T \mathbf{B})} \quad (114)$$

**B** ( $P, C$ )-Matrix der partiellen Empfindlichkeiten

Aus einem überbestimmten System von  $P$  Messstellen (Sensoren, Wellenlängen, ...) könnte man also schrittweise zum Beispiel diejenigen  $P_{OPT} \geq C$  Messstellen auswählen, für die gilt:

$$sen_{OPT} = \max_S (sen_s) \quad (115)$$

Darin bedeutet  $S$  die Anzahl der möglichen Messstellenkombinationen, die nach

$$S = \binom{P}{C} = \frac{P!}{C! (P - C)!} \quad (116)$$

berechnet werden kann.<sup>63</sup>

❶ Datenbeispiel siehe Abschnitt 5.7.2

### 5.3.2 ... NACH DER SELEKTIVITÄT (MESSSTELLEN-KRITERIUM)

Für den Fall  $P = C$  hat KAISER den Begriff *Selektivität* ( $sel$ ) eingeführt. Dabei stellt man die partielle Empfindlichkeit einer Messstelle (einer Wellenlänge, „eines Sensors“) für einen Analyten den partiellen Empfindlichkeiten der anderen Komponenten („Störkomponenten“) gegenüber:

$$sel_p = \frac{|b_{pp}|}{\sum_{c=1; c \neq p}^C |b_{pc}|} - 1 \quad (117)$$

62 A. JUNKER, G. BERGMANN: „Selection, comparison and validation of optimum working conditions for qualitative multi-component analysis. Part I. Two-dimensional overdetermined systems, Sensitivity as optimisation parameter“, Z. Anal. Chem. 272(1974)267; zitiert nach MASSART et al. (1988, S. 126)

63 Studenten wird empfohlen, bei dieser Gelegenheit den Unterschied zwischen Permutation, Variation und Kombination von verschiedenen Elementen mit und ohne Wiederholung nachzulesen. Für  $C = 2$  ergibt sich aus dem quadratischen Schema der Wellenlängen-Kombinationen sofort  $S = P(P - 1) / 2$ .

Im günstigsten Fall ist eine solche partielle Selektivität also unendlich groß. Das Sensorsystem wird jedoch durch die geringste Selektivität bestimmt:

$$sel_{SYSTEM} = \min_P (sel_p) \quad (118)$$

Man wird also dasjenige System von  $P$  Messstellen oder, bei überbestimmten Systemen, diejenige Untermenge  $P_{OPT} \geq C$  der Messstellen auswählen, deren minimale Selektivität am größten ist.

$$sel_{OPT} = \max_S (sel_{SYSTEM}) \quad (119)$$

In einem voll selektiven Verfahren sind nur die  $b_{pp} \neq 0$ .

Mit einem völlig selektiven Analysenverfahren ist also eine unabhängige Bestimmung aller Komponenten nebeneinander möglich.

(KAISER spricht davon, dass ein voll selektives Verfahren für alle Komponenten partiell spezifisch ist; siehe unten.)

In einem Zweikomponentensystem sind die Verhältnisse besonders überschaubar.

Jedoch wird auch hier kritisiert, dass eine unabhängige Bewertung der Wellenlängen erfolgt, weshalb ein „Selektivitätsprodukt“<sup>64</sup> vorgeschlagen wurde:

$$sel_{PRO} = \frac{b_{11} b_{22}}{b_{12} b_{21}} \quad (120)$$

❶ Datenbeispiel siehe Abschnitt 5.7.3

### 5.3.3 ... NACH DER SPEZIFITÄT (KOMPONENTEN-KRITERIUM)

Vergleicht man die partielle Empfindlichkeit einer Messstelle (Sensor, Wellenlänge, ...) bezüglich einer Komponente mit allen anderen möglichen Analysenkomponenten, so erhält man nach KAISER den *Grad der Spezifität* aus

$$spe_p = \frac{|b_{pp}|}{\sum_{c=1; c \neq p}^C |b_{cc}|} - 1 \quad (121)$$

64 S. EBEL, E. GLASER, SAA ABDULLA, U. STEFFENS, V. WALTER: „Auswahl optimaler Messwellenlängen bei der spektroskopischen Mehrkomponenten-Analyse. 1. Bestimmtes Gleichungssystem“, Fresenius Z. Anal. Chem. 313(1982)24-27

Für ein System aus  $P$  Messstellen (Sensoren, ...) kann als gemeinsame Spezifität wieder der ungünstigste Fall angenommen werden:

$$spe_{SYSTEM} = \min_P (spe_p) \quad (122)$$

In einem spezifischen Verfahren ist nur ein einziges  $b_{pp} \neq 0$ , d. h. mit einem solchen Verfahren kann überhaupt nur eine Komponente bestimmt werden, da alle anderen Komponenten keinen Signalbeitrag liefern.

Im Falle überbestimmter Systeme (mehrere mögliche Messstellen, Sensoren etc.) wird man wie oben diejenige optimale Sensorkombination auswählen, die die größte minimale Spezifität hat:

$$spe_{OPT} = \max_S (spe_{SYSTEM}) \quad (123)$$

( $S$  hat die gleiche Bedeutung wie oben.)

❶ Datenbeispiel siehe Abschnitt 5.7.3

#### 5.3.4 ... MITTELS ANDERER METHODEN

Die genannten KAISERschen Kriterien sind aus der Sicht moderner Mehrkomponenten-Analytik schon von OTTO und WEGSCHNEIDER (1986) kritisch beurteilt worden. Zur Bewertung von Kalibriermodellen gibt es zahlreiche andere Vorgehensweisen. Ein Methodenentwickler wird die für seine analytische Aufgabe am besten geeignete Variante auswählen.

Wir nennen hier exemplarisch die ersten beiden Diplomarbeiten aus unserer Merseburger Arbeitsgruppe.

GRÄHLERT<sup>65</sup> hat sich speziell mit „Optimaler Messstellenauswahl in der photometrischen Mehrkomponentenanalytik“ befasst und folgende Aspekte erfasst:

- (1) Beurteilung einzelner Wellenlängen mittels
  - Varianz
  - Distension
  - informationstheoretischer Kriterien

---

65 W. GRÄHLERT: „Optimale Messstellenauswahl in der photometrischen Mehrkomponentenanalytik“, Diplomarbeit, Merseburg, 1990

- (2) Beurteilung von Wellenlängen-Mengen mittels
- Determinante
  - Konditionszahl
  - Distension
  - informationstheoretischer Kriterien
  - Ähnlichkeitskriterium
- (3) Als Suchkriterien für die Teilmengen untersuchte er
- Optimierungsverfahren (stochastisches Verfahren; Polytopverfahren)
  - Screeningverfahren

Über „Chemometrische Methoden zur Beurteilung und Auswahl von Kalibrierexperimenten“ hat KRAUSE<sup>66</sup> berichtet. Sein Thema war die simultane Mehrkomponenten-Analyse von Ni, Zn und Pb. Zum Kalibrieren benutzte er die Methode PLS und zur Auswahl der Kalibrierproben zog er die Cluster- und Diskriminanzanalyse heran.

## 5.4 Anwenden des Kalibrationsmodells

Das Kalibrationsmodell wird sowohl während der Validierphase („Testphase“) als auch zur Bestimmung unbekannter Analytwerte (Konzentrationen) in Analyseproben angewendet (siehe Gleichung 111).

Im Analysenschritt, beim „Analysieren“, gibt es zwei Fälle, um von den gemessenen Signalwerten, die im Messwertvektor  $\mathbf{y}$  zusammengefasst sind, zu den gesuchten Analytwerten zu kommen, die im Vektor  $\mathbf{x}$  zusammengefasst sind.

- (1) Bei  $P = C$ , d. h. exakt bestimmten Kalibrier- bzw. Gleichungssystemen, ist die Empfindlichkeitenmatrix eine quadratische Matrix, die direkt invertiert werden kann:

$$\hat{\mathbf{x}} = \hat{\mathbf{B}}^{-1} \mathbf{y} \quad (124)$$

- (2) Bei  $P > C$ , also überbestimmten Kalibrier- bzw. Gleichungssystemen, ergibt sich der Analytwertevektor aus:

$$\hat{\mathbf{x}} = (\hat{\mathbf{B}}^T \hat{\mathbf{B}})^{-1} \hat{\mathbf{B}}^T \mathbf{y} \quad (125)$$

Es ist also in jedem Fall eine Matrixinversion auszuführen.

Formulierungen mit Matrizen wie vor dem Vektor  $\mathbf{y}$  werden in der linearen Algebra bzw. numerischen Mathematik *generalisierte* oder *Pseudoinverse* der jeweiligen Matrix genannt. Deswegen heißen die Methoden unter Verwendung von Gleichung (125) im englischen Schrifttum *generalized inverse methods*.

Bei der praktischen Ausführung solcher Inversionen, der Berechnung der inversen Empfindlichkeitenmatrix bzw. der *Pseudoinversen* der Konzentrations- oder Empfindlichkei-

---

66 S. KRAUSE: „Chemometrische Methoden zur Beurteilung und Auswahl von Kalibrierexperimenten“, Diplomarbeit, Merseburg, 1990

tenmatrix können numerische Probleme auftreten, die die Zuverlässigkeit der Ergebnisse beeinträchtigen können.<sup>67</sup>

Im Gegensatz zu den beiden Varianten (1) und (2), die man auch unter dem Begriff *K-Matrix-Methoden* zusammenfasst, geht man deshalb bei der *P-Matrix-Methode*<sup>68</sup> direkt von dem Anwendungsmodell aus:

$$\mathbf{x} = \mathbf{P} \mathbf{y} \quad (126)$$

Offensichtlich ist dies auch eine Modellgleichung der inversen Kalibration. Im Kalibrierschritt wird die *Vorhersagematrix*  $\mathbf{P}$  (engl.: prediction) dann unmittelbar geschätzt, und man umgeht im Anwendungsschritt die Inversion:

$$\hat{\mathbf{x}} = \hat{\mathbf{P}} \mathbf{y} \quad (127)$$

Mitunter können die numerischen Probleme bei der Berechnung von Pseudoinversen der Konzentrations- bzw. der Signalmatrix dadurch umgangen werden, dass eine mathematische *Hauptkomponentenzerlegung* als Datenvorbereitungsschritt vorgenommen wird. Damit sichert man, dass die ehemals linear abhängigen Zeilen bzw. Spalten der Rohdatenmatrizen in orthogonale Zeilen bzw. Spalten umgewandelt werden. Dasselbe Abhängigkeitsproblem verbirgt sich auch hinter dem Begriff der *Multikollinearität* in der üblichen multivariaten Regressionsanalyse.

❶ Datenbeispiel siehe Abschnitt 5.7.4

## 5.5 Kenngrößen für Fehler der Kalibration

Auch in der „multivariaten Welt“ gelten prinzipiell die Definitionen, die wir in Teil A kennen gelernt haben. Aktuelle Geräte- bzw. Firmensoftware wird mit den gleichen Begriffen arbeiten.

Wir beschränken uns hier auf den Hinweis:

Grundlage ist auch im multivariaten Kalibrierfall die Kenngröße *PRESS* (predictive sum of squares), die Summe der quadrierten Differenzen aus Vorhersage- und Referenzwerten.

Aus ihr leitet sich wiederum der Kalibrierfehler *SEC* (standard error of calibration, Modellfehler) bzw. der Validier- oder Vorhersagefehler *SEP* (standard error of prediction) ab, der von guter Software nach Kreuzvalidierung als *SEPCV* ausgewiesen wird (CV = cross validation). Bei der Kreuzvalidierung werden nach bestimmten Verfahren nacheinander Untermengen der Validierproben ausgewählt. Mit diesen werden dann die jeweiligen Kenngrößen berechnet, woraus schließlich ein „realistischerer“ Validierfehler ermittelt wird (vgl. Abschnitt 2.8).

67 Im Zusammenhang mit quantenchemischen Rechnungen, bei denen hochdimensionale Matrizen auftreten können, wurde bereits darauf hingewiesen, dass die Auflösung eines linearen Gleichungssystems numerisch nicht genauer erfolgen kann, als deren Elemente durch Ziffern dargestellt werden können. (Siehe: H. ZWANZIGER, J. REINHOLD: „Zu einigen numerischen Problemen bei quantenchemischen Rechnungen“, *Wiss. Z. Universität Leipzig* 28(1979)655-659)

68 C. W. BROWN, P. F. LYNCH, R. J. OBREMSKI, D. S. LAVERY: „Matrix representations and criteria for selecting analytical wavelengths for multicomponent spectroscopic analysis“, *Anal. Chem.* 54(1982)1472-1479

## 5.6 Maximalfehler des Ergebnisses

Die besprochenen KAISERSchen Definitionen enthalten explizit weder den Fehler der (Signal-)Messung noch den der zu erwartenden Analyt-Ergebnisse (Konzentrationen). Implizit steckt der Fehler der Signalmessungen natürlich in den partiellen Empfindlichkeiten.

Da die zu erwartende Präzision der Ergebnisse häufig ein Optimierungskriterium ist, muss auch der relative Fehler der Analytwerte zur Auswahl des Kalibriersystems herangezogen werden.

Die numerische Mathematik<sup>69</sup> liefert dazu vereinfacht als erwartete obere Grenze der relativen Abweichung eines Ergebnisses:

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \|\mathbf{B}\| \|\mathbf{B}^{-1}\| \left( \frac{\|\Delta \mathbf{y}\|}{\|\mathbf{y}\|} + \frac{\|\Delta \mathbf{B}\|}{\|\mathbf{B}\|} \right) \quad (128)$$

$\Delta \mathbf{x}, \Delta \mathbf{y}, \Delta \mathbf{B}$	absolute Abweichungen (Differenzen)
$\ \dots\ $	geeignete Norm der Abweichungen (Definitionen im folgenden Exkurs)
$\mathbf{B}$	Empfindlichkeitenmatrix

Das Produkt vor der Klammer wird als *relative Konditionszahl*  $rcond(\mathbf{B})$  bezeichnet.

Sie wäre ein direktes Maß für die Fehlerverstärkung der relativen Messwertabweichung, wenn man den „Fehler“  $\|\Delta \mathbf{B}\|/\|\mathbf{B}\|$  bei der Bestimmung der Empfindlichkeitenmatrix  $\mathbf{B}$  vernachlässigen kann.

Die Norm  $\|\mathbf{B}^{-1}\|$  bezeichnet man auch als absolute Konditionszahl.

Die beste Konditionszahl ist gleich Eins, und man wird deshalb diejenigen  $P = C$  Messstellen (Sensoren, Wellenlängen, ...) auswählen, für die  $rcond(\mathbf{B})$  möglichst klein ausfällt.

Die Konditionszahl ist deswegen besonders geeignet, weil sie berücksichtigt, dass bei der Auflösung unseres Gleichungssystems von Kalibriergleichungen neben  $\mathbf{B}$  auch die Inverse von  $\mathbf{B}$  verwendet wird.

❶ Datenbeispiel siehe Abschnitt 5.7.5

### Exkurs: Vektor- und Matrixnormen

Dieser Abschnitt ist insbesondere für die mathematisch Interessierten bestimmt, die ggf. selbst die oben eingeführten Maßzahlen programmieren wollen oder müssen.

Meist wird als Vektornorm  $\|\mathbf{x}\|$  die EUKLIDische Norm verwendet:

---

69 G. MAESZ: „Vorlesungen über numerische Mathematik I“, Akademie-Verlag, Berlin, 1984, S. 99ff.

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{n=1}^N x_n^2} = \sqrt{\mathbf{x}^T \mathbf{x}} \quad (129)$$

Die dieser Vektornorm zugeordnete Matrixnorm  $\|\mathbf{A}\|$  erhält man nach Lösung des speziellen Eigenwertproblems:

$$\det(\mathbf{A} - \eta \mathbf{I}) = 0 \quad (130)$$

wobei  $\mathbf{A}$  eine quadratische  $(N,N)$ -Matrix,  $\mathbf{I}$  die diagonale Einheitsmatrix und  $\eta$  ein Eigenwert von  $\mathbf{A}$  ist.

Mit dem Spektralradius von  $\mathbf{A}$ , dem maximalen Eigenwert  $\eta_{MAX}$  von  $\mathbf{A}$ , bzw. dem kleinsten Eigenwert  $\eta_{MIN}$

$$\begin{aligned} \eta_{MAX}(\mathbf{A}) &= \max_N (|\eta_n|) \\ \eta_{MIN}(\mathbf{A}) &= \min_N (|\eta_n|) \end{aligned} \quad (131)$$

sind folgende Spektralnormen von  $\mathbf{A}$  definiert:

$$\|\mathbf{A}\|_2 = \sqrt{\eta_{MAX}(\mathbf{A}^T \mathbf{A})} \quad \|\mathbf{A}^{-1}\|_2 = \frac{1}{\sqrt{\eta_{MIN}(\mathbf{A}^T \mathbf{A})}} \quad (132)$$

Für quadratische *und* symmetrische Matrizen  $\mathbf{A}$  gilt vereinfacht:

$$\|\mathbf{A}\|_2 = \eta_{MAX}(\mathbf{A}) \quad \|\mathbf{A}^{-1}\|_2 = \frac{1}{\eta_{MIN}(\mathbf{A})} \quad (133)$$

Für die quadratische Empfindlichkeitenmatrix  $\mathbf{B}$  eines exakt bestimmten Kalibriersystems mit  $P = C$  ergibt sich folgende relative Konditionszahl:

$$rcond(\mathbf{B}) = \frac{\eta_{MAX}(\mathbf{B})}{\eta_{MIN}(\mathbf{B})} \quad (134)$$

Die Berechnung der Spektralnormen kann aufwändig sein, weswegen man als passende Normen zur EUKLIDischen Vektornorm auch die FROBENIUS- (oder SCHUR-)Norm oder die *Maximumnorm* verwenden kann:

$$\|\mathbf{B}\|_{FROB} = \sqrt{\sum_{p=1}^P \sum_{c=1}^C b_{pc}^2} \quad (135)$$

$$\|\mathbf{B}\|_{MAX} = N \max (b_{pc}) \quad (136)$$

Anstelle der Maximumnorm kann ebenso einfach die Zeilensummennorm  $\|\mathbf{B}\|_{\infty}$  oder die Spaltensummennorm  $\|\mathbf{B}\|_1$  berechnet werden.

Allerdings kann es weiterhin numerische Probleme beim Invertieren von  $\mathbf{B}$  geben.

❶ Datenbeispiel siehe Abschnitt 5.7.5

## 5.7 Datenbeispiele<sup>70</sup>

### 5.7.1 DATENBEISPIEL - VISUELLE SPEKTRENBEWERTUNG

Wie wichtig die visuelle Inspektion von Datensätzen ist, haben wir bereits mehrfach betont. Das gilt gleichermaßen für Spektren-Darstellungen.

❶ Daten-ID: 15.01.2020; Spektrum aufgenommen mit GENESYS 180, Thermo Fisher Scientific

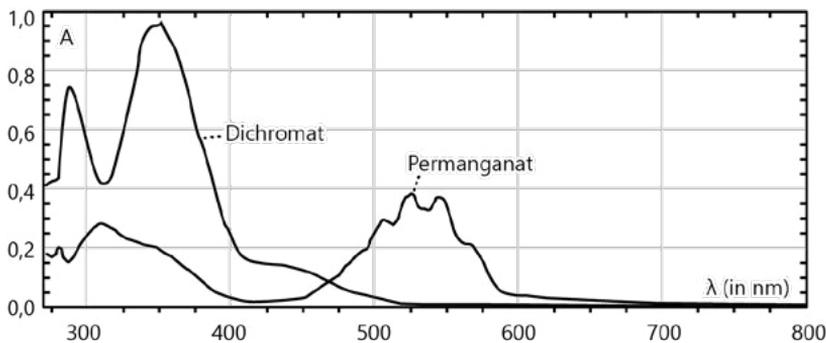


Abb. 12: UV-Spektren von Permanganat und Dichromat

Die nachfolgend verwendeten Daten (Liane Aue, Dezember 2019) beruhen auf einem Praktikumsversuch (Hochschule Merseburg) zur simultanen fotometrischen Bestimmung von  $\text{MnO}_4^-$  und  $\text{Cr}_2\text{O}_7^{2-}$ . Die Analyte werden im Folgenden mit Mn bzw. Cr abgekürzt.

Bei  $\lambda_1 = 526 \text{ nm}$  scheint die empfindlichste Messstelle von Mn (in Permanganatlösungen) zu liegen; die Absorbanz von Cr (in Dichromatlösungen) ist bei dieser Wellenlänge zu vernachlässigen. Eine ungestörte Bestimmung von Mn neben Cr sollte also möglich sein. Bei  $\lambda_2 = 350 \text{ nm}$ , der empfindlichsten Messstelle von Cr, ist jedoch die Absorbanz durch Mn noch merklich. Eine Bestimmung von Cr neben Mn bei dieser Wellenlänge wird also durch die Querempfindlichkeit gestört und wahrscheinlich systematisch verfälscht.

70 In der Regel werden alle unabhängigen, gegebenen, übernommenen (x-)Werte kursiv gesetzt.

Bei allen Datenanalysen sollten stets auch visuelle Darstellungen genutzt werden. Sie können numerische Ergebnisse bildhaft ergänzen und dadurch manchmal besser vermitteln. In der Spektroskopie sind Spektren ohnehin von zentraler Bedeutung.

## 5.7.2 DATENBEISPIEL ZU 5.2 - MODELLIERUNG, MODELLBEWERTUNG

Modellierung und Modellbewertung bilden natürlich eine Einheit. Der Aufstellung eines Kalibriermodells folgt die Anwendung einer geeigneten Regressionsrechnung, auf deren statistisches Modell wir nicht eingehen müssen. Die Bewertung ihrer Ergebnisse, in erster Linie der Schätzwerte für die Empfindlichkeiten ist zugleich eine Bewertung unseres jeweiligen Kalibriermodells.

Wir überlassen es den Nutzern, die folgenden Ergebnisse mit ihrer bevorzugten Software nachzuvollziehen. Wir nutzen in der Regel das Programmpaket STATISTICA® (Statsoft Inc.) oder die Möglichkeiten von Excel® (Microsoft Corporation).

❶ Daten-ID: 19.12.2019; M-Proben, Design: 2<sup>2</sup>-VP mit Zentralpunkt; Mittelwerte aus Dreifachmessungen bei zwei Wellenlängen

In fünf Mischungen beider Analyte wurden die folgenden Absorbanzen gemessen. Alle Konzentrationen sind in mmol l<sup>-1</sup> angegeben. Ihre Werte folgen einem optimalen Versuchsplan mit Zentralpunkt.

Die Einheiten der Größen  $X$  und  $Y$  sind also:  $[X] = \text{mmol l}^{-1}$ ,  $[Y] = \text{ohne}$  (Die Absorbanz oder Extinktion ist eine „dimensionslose“ Größe.),  $[\lambda] = \text{nm}$ .

Zur Zahlengenauigkeit von Rechenergebnissen siehe Fußnote.<sup>71</sup>

Probe	Sollkonzentration		Absorbanz bei	
	Mn	Cr	526 nm	350 nm
M1.1	0,08	0,15	0,186	0,583
M1.2	0,08	0,15	0,186	0,583
M1.3	0,08	0,15	0,185	0,583
M2.1	0,08	0,45	0,191	1,553
M2.2	0,08	0,45	0,192	1,559
M2.3	0,08	0,45	0,192	1,559
M3.1	0,32	0,15	0,764	0,884
M3.2	0,32	0,15	0,764	0,885
M3.3	0,32	0,15	0,764	0,885
M4.1	0,32	0,45	0,770	1,831
M4.2	0,32	0,45	0,770	1,830
M4.3	0,32	0,45	0,770	1,831
M5.1	0,20	0,30	0,478	1,262
M5.2	0,20	0,30	0,478	1,262
M5.3	0,20	0,30	0,478	1,263

### (1) „Zweipunkt“-Modellierung

Mit den Messwerten der jeweils ersten vier Kalibriermischungen ergeben sich mittels *multipler Regression* für die empfindlichsten Wellenlängen die Koeffizienten der *unabhängigen Modellgleichungen*

<sup>71</sup> Unsere Rechenwerte sind (willkürlich) meist auf drei Nachkommastellen gerundet. Damit sind sie ausreichend „genau“, um miteinander verglichen werden zu können, auch wenn sie manchmal „so genau“ gar nicht relevant sind. Ergebnisse aus Rechenprogrammen werden häufig mit einer völlig sinnlosen Genauigkeit angegeben. Der deutsche Mathematiker, Astronom und Physiker Carl Friedrich GAUSS (1777-1855) hat dazu einmal bemerkt, dass sich „der Mangel an mathematischer Bildung“ nirgends mehr als in einer übertrieben genauen Zahlendarstellung zeigt. (Vgl. das Motto in F. W. KÜSTER, A. THIEL: „Analytik. Daten, Formeln, Übungsaufgaben“, de Gruyter, Berlin, Boston, 108. Aufl., 2016)

$$Y(\lambda_{1,MAX}) = f(Mn, Cr) = b_{10} + b_{11} X_{1(=Mn)} + b_{12} X_{2(=Cr)}$$

und

$$Y(\lambda_{2,MAX}) = f(Mn, Cr) = b_{20} + b_{21} X_{1(=Mn)} + b_{22} X_{2(=Cr)}$$

in folgender Tabelle:

Modelltyp	Leerwert		Empfindlichkeit (Mn)		Empfindlichkeit (Cr)	
$Y_{(\lambda_{MAX,1})} = f(Mn, Cr)$	$b_{10} =$	-0,010	$b_{11} =$	2,410	$b_{12} =$	0,020
$Y_{(\lambda_{MAX,2})} = f(Mn, Cr)$	$b_{20} =$	0,014	$b_{21} =$	1,199	$b_{22} =$	3,200

Die errechneten Empfindlichkeiten bestätigen die Aussagen der Spektrenbewertung. Die Modellkoeffizienten sind allesamt signifikant und ihre Fehler sind sehr gering. (Die Determinante dieses Kalibriersystems beträgt 7,7. Sie hat aber jetzt noch keine Bedeutung.)

Bezüglich der Form der Kalibriergraphen haben wir es jetzt mit Flächen zu tun, die für jede Wellenlänge die Abhängigkeit ihrer Messwerte von den beiden Konzentrationsgrößen darstellen.

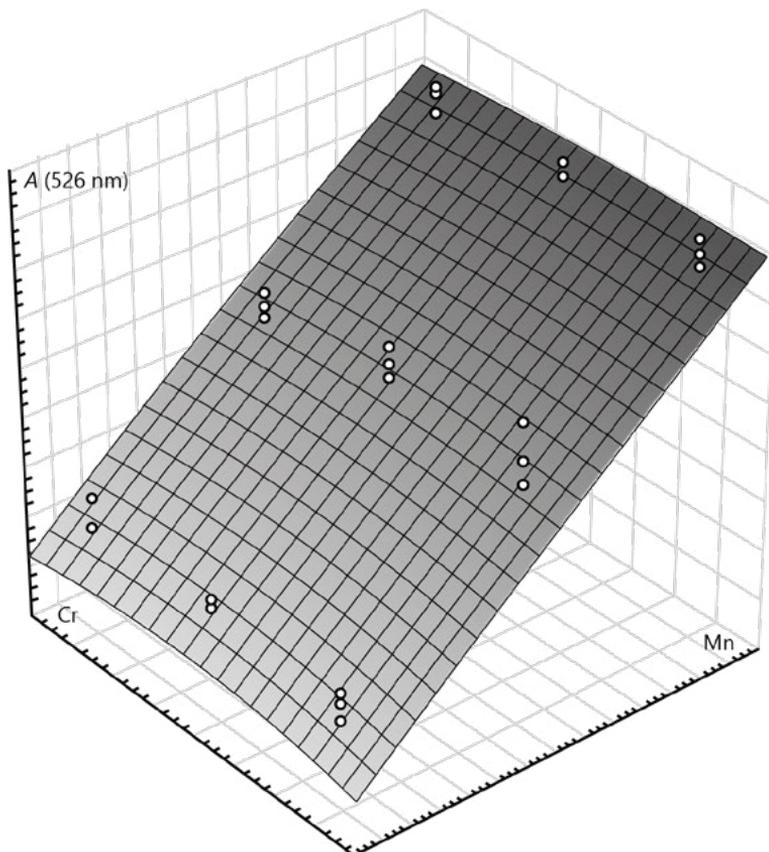


Abb. 13: Univariate Kalibrierfläche  $Y = f(Mn, Cr)$

Auch die Modell-Bewertung mittels der beiden signalbezogenen Kenngrößen *Bestimmtheitsmaß* ( $B$ ;  $r^2$ ) und *Regressionsfehler* ( $s_{RES}$ ) fällt positiv aus.

Diese „guten“ Werte sind nicht allzu verwunderlich, denn das Proben-Design bildet bisher, bezüglich jeden einzelnen Analyten gesprochen, eine Art „Zweipunkt-Kalibration“ ab, und durch zwei Punkte verlaufen eben die „besten“ Geraden. Im nächsten Beispiel-Kästchen machen wir das Proben-Design deshalb etwas realistischer.

Fehler der Modellparameter	Leerwert		Empfindlichkeit (Mn)		Empfindlichkeit (Cr)		$B$	$s_{RES}$	$SEC$	$SEP$
Modell $Y(\lambda_1)$	$s(b_{10}) =$	0,000	$s(b_{11}) =$	0,001	$s(b_{12}) =$	0,001	1,000	0,000		
Modell $Y(\lambda_2)$	$s(b_{20}) =$	0,007	$s(b_{21}) =$	0,020	$s(b_{22}) =$	0,016	1,000	0,008		

## (2) „Dreipunkt“-Modellierung

Wir gehen nun zu einem realistischeren Proben-Design über, indem wir unserem immerhin vollständigen zweifaktoriellen „2<sup>2</sup>-Versuchsplan“ noch einen „Zentrumspunkt“ hinzufügen. Damit nähern wir uns einem zweifaktoriellen, dreistufigen „3<sup>2</sup>-Versuchsplan“ und können quasi eine Art „Dreipunkt-Kalibration“ abbilden.

Mit den hinzukommenden Daten des Zentralpunkts ändert sich an den bisherigen Ergebnissen nicht allzu viel. Für Mn bleiben die Koeffizienten und Fehler unverändert. Bei Cr ändert sich der Leerwert, das Bestimmtheitsmaß wurde etwas kleiner, der Regressionsfehler etwas größer. (Die Determinante ist ebenfalls unverändert.)

Modellparameter (Koeffizienten)	Leerwert		Empfindlichkeit (Mn)		Empfindlichkeit (Cr)	
$Y(\lambda_1) = f(Mn, Cr)$	$b_{10} =$	-0,010	$b_{11} =$	2,410	$b_{12} =$	0,020
$Y(\lambda_2) = f(Mn, Cr)$	$b_{20} =$	0,024	$b_{21} =$	1,199	$b_{22} =$	3,200

Fehler der Modellparameter	Leerwert		Empfindlichkeit (Mn)		Empfindlichkeit (Cr)		$B$	$s_{RES}$	$SEC$	$SEP$
Modell $Y(\lambda_1)$	$s(b_{10}) =$	0,000	$s(b_{11}) =$	0,001	$s(b_{12}) =$	0,001	1,000	0,000	0,000	0,002
Modell $Y(\lambda_2)$	$s(b_{20}) =$	0,018	$s(b_{21}) =$	0,055	$s(b_{22}) =$	0,044	0,998	0,023	0,008	0,006

Von den beiden *konzentrationsbezogenen Kenngrößen*, Kalibrierfehler ( $SEC$ ) und Vorhersage- oder Validierfehler ( $SEP$ ), kann  $SEC$  anhand der Kalibrierproben unmittelbar berechnet werden.

Die Daten zur Berechnung des Vorhersagefehlers anhand von Validierproben, deren Messwerte nicht zur Modellierung verwendet wurden, werden im Beispiel-Abschnitt zur Validierung angegeben.

### 5.7.3 DATENBEISPIEL ZU 5.3 - KAISERSCHE AUSWAHLKRITERIEN

Schon frühzeitig hat man sich Gedanken gemacht, wie man am besten die „Sensitivität“ oder „Spezifität“ einer Analysenmethode quantifizieren kann. Von KAISER stammen systematische Ansätze dazu, die wir im folgenden Abschnitt nachvollziehen wollen.

❶ Daten-ID: 19.12.2019; Mischungsproben, Design: 3<sup>2</sup>-VP; Mittelwerte aus Dreifachmessungen bei zwei Wellenlängen

Wir verwenden die Mittelwerte von Proben mit jeweils den beiden extremen Konzentrationen, die für unsere „einfachen Anstiegsberechnungen“ ausreichen. Aus den Messwerten bei 350 nm und 526 nm wurden mittels der Verhältnisse in den beiden Spektren die Werte für 290 nm und 544 nm errechnet:

Probe	Konzentration		Absorbanz bei			
	Mn	Cr	290 $\lambda_1$	350 $\lambda_2$	526 $\lambda_3$	544 nm $\lambda_4$
M11	0,08	0,15	0,454	<b>0,601</b>	<b>0,193</b>	0,186
M31	0,32	0,15	0,685	0,907	0,761	0,732
M13	0,08	0,45	1,163	1,540	0,196	0,189
M33	0,32	0,45	1,381	1,828	0,767	0,738

Mittels multipler Regression ermitteln wir für jede Wellenlänge ein Kalibriermodell. Die partiellen Empfindlichkeiten  $b$  in den jeweiligen Modellen sind:

bei	$b(\text{Cr})$	$b(\text{Mn})$	Leerwert
$\lambda_1$	2,341	0,935	0,031
$\lambda_2$	3,100	1,238	0,042
$\lambda_3$	0,016	2,374	0,000
$\lambda_4$	0,015	2,283	0,000

Mit diesen Empfindlichkeiten ergeben sich für die möglichen 5 Wellenlängen-Kombinationen die Werte der besprochenen Kriterien:

	$sen, \det(\mathbf{B})$	min $sel_s$	min $spe_s$	$sel_{PRO}$
$\lambda_1, \lambda_2$	0	-0,6	-0,47	1
$\lambda_1, \lambda_3$	5,5	<b>1,5</b>	<b>-0,01</b>	<b>382</b>
$\lambda_1, \lambda_4$	5,3	<b>1,5</b>	-0,02	<b>382</b>
$\lambda_2, \lambda_3$	<b>7,3</b>	<b>1,5</b>	-0,23	<b>382</b>
$\lambda_2, \lambda_4$	7,1	<b>1,5</b>	-0,26	<b>382</b>
$\lambda_3, \lambda_4$	0	-1	-0,99	1

Fazit:

- (1) Das von uns aus numerischen Gründen bevorzugte Determinanten-Kriterium bestätigt das bisher verwendete Kalibrationsmodell mit  $\{\lambda_2, \lambda_3\}$ .
- (2) Das Spezifitätskriterium bevorzugt die Wellenlängen-Kombination  $\{\lambda_1, \lambda_3\}$ . Dort hat die Empfindlichkeitenmatrix tatsächlich die „größte“ Diagonal-Dominanz.
- (3) Das KAISERSche Selektivitäts-Kriterium und das „Selektivitätsprodukt“ lassen auch noch andere Kombinationen zu.

- (4) Nicht unerwartet: Wellenlängen(bereiche), wo hauptsächlich Mn- oder Cr-Species absorbieren, werden bei allen Varianten ausgeschlossen.

#### 5.7.4 DATENBEISPIEL ZU 5.4 - ANWENDEN DES KALIBRIERMODELLS

Das Aufstellen eines Kalibriermodells und seine statistische Bewertung sind wichtige Vorarbeiten. Aber nicht die „Signalmodellierung“ ist das primäre Ziel des Analytikers. Er will das Modell anwenden können, um Analytwerte zu bestimmen. Die erste Anwendung geschieht bereits im Validierungsschritt (Teil 2 dieses Abschnitts), der deshalb auch als Modellprüfung angesehen werden kann, während die eigentliche Anwendung, der „Praxistest“, im Analysenschritt erfolgt (Teil 3 dieses Abschnitts). Für beide Anwendungen muss das Kalibriermodell „invertiert“ werden. Damit befasst sich Teil (1) dieses Abschnitts.

- ❶ Daten-ID: Modellparameter aus Messungen vom 19.12.2019; Werte wie oben; Validierung mit Mischungsproben vom 19.12.2019; A-Proben vom 19.12.2019

##### (1) Invertieren des Kalibriermodells

Bei der Anwendung des gefundenen Kalibriermodells auf eine Analysenprobe muss, wie schon zur Berechnung der beiden *konzentrationsbezogenen Kenngrößen* Kalibrierfehler (*SEC*) und Vorhersagefehler (*SEP*), die jeweilige Modellgleichung für Mn bzw. Cr nach der entsprechenden Konzentrationsgröße umgestellt werden.

Da die Schätzwerte (Rechenwerte) für die Modellparameter bekannt sind, ist dies ohne Probleme möglich. Eigentlich sollte der Fall  $P = C$  vorliegen, allerdings haben wir durch die Leerwerte keine quadratische, sondern eine (2,3)-Matrix zu invertieren.

Wir können das aber umgehen, indem wir die gemessenen Signalwerte um den Leerwert korrigieren. Danach könnten wir mit der quadratischen Matrix der Empfindlichkeiten operieren.

Dass dies ein legitimer Schritt ist, zeigt die Auflösung des Gleichungssystems:

$$X_1 = \frac{b_{22}(Y_1 - b_{10}) - b_{12}(Y_2 - b_{20})}{b_{11}b_{22} - b_{12}b_{21}}$$

$$X_2 = \frac{(Y_2 - b_{20}) - b_{21}X_1}{b_{22}}$$

Die Klammerausdrücke sind die leerwertkorrigierten Signalgrößen.

Gleichzeitig erkennt man im Nenner die Determinante des Kalibriersystems, womit auch deren Auswirkung bei der Matrixinversion klar wird, sobald sie gegen Null geht.

##### (2) Validieren des Kalibriermodells

Zur Validierung werden Proben verwendet, die nach einem  $3^2$ -Versuchsplan hergestellt wurden.

Proben	Konz.		Absorbanz bei (nm)		berechn. Konz.		Differenz	
	Mn	Cr	526	350	Mn	Cr	Mn	Cr
M11.1	0,08	0,15	0,193	0,601	0,08	0,15	0,00	0,00
M11.2	0,08	0,15	0,193	0,601	0,08	0,15	0,00	0,00
M11.3	0,08	0,15	0,193	0,601	0,08	0,15	0,00	0,00
M12.1	0,08	0,30	0,193	1,062	0,08	0,30	0,00	0,00
M12.2	0,08	0,30	0,193	1,062	0,08	0,30	0,00	0,00
M12.3	0,08	0,30	0,191	1,063	0,08	0,30	0,00	0,00
M13.1	0,08	0,45	0,196	1,540	0,08	0,45	0,00	0,00
M13.2	0,08	0,45	0,196	1,540	0,08	0,45	0,00	0,00
M13.3	0,08	0,45	0,196	1,540	0,08	0,45	0,00	0,00
M21.1	0,20	0,15	0,478	0,768	0,20	0,16	0,00	0,01
M21.2	0,20	0,15	0,478	0,767	0,20	0,16	0,00	0,01
M21.3	0,20	0,15	0,479	0,770	0,20	0,16	0,00	0,01
M22.1	0,20	0,30	0,479	1,210	0,20	0,30	0,00	0,00
M22.2	0,20	0,30	0,479	1,210	0,20	0,30	0,00	0,00
M22.3	0,20	0,30	0,481	1,211	0,20	0,30	0,00	0,00
M23.1	0,20	0,45	0,481	1,679	0,20	0,45	0,00	0,00
M23.2	0,20	0,45	0,481	1,679	0,20	0,45	0,00	0,00
M23.3	0,20	0,45	0,481	1,679	0,20	0,45	0,00	0,00
M31.1	0,32	0,15	0,761	0,907	0,32	0,16	0,00	0,01
M31.2	0,32	0,15	0,761	0,907	0,32	0,16	0,00	0,01
M31.3	0,32	0,15	0,761	0,907	0,32	0,16	0,00	0,01
M32.1	0,32	0,30	0,764	1,342	0,32	0,30	0,00	0,00
M32.2	0,32	0,30	0,764	1,342	0,32	0,30	0,00	0,00
M32.3	0,32	0,30	0,761	1,334	0,32	0,29	0,00	-0,01
M33.1	0,32	0,45	0,770	1,830	0,32	0,45	0,00	0,00
M33.2	0,32	0,45	0,766	1,827	0,32	0,45	0,00	0,00
M33.3	0,32	0,45	0,766	1,827	0,32	0,45	0,00	0,00

Mittels der Koeffizienten des oben dargestellten „Dreipunkt-Kalibrationsmodells“ wurden aus den Absorbanzen die Konzentrationen für Mn und Cr berechnet und über die Kenngröße *PRESS* die oben tabellierten *SEP*-Werte berechnet.

### (3) Analytische Anwendung des Kalibriermodells

Anhand zweier Proben, deren Werte bekannt sind, testen wir unsere Modelle und erhalten ein akzeptables Ergebnis.

Probe	Konzentration		Absorbanz bei $\lambda$ (in nm)		Berechn. Konz.		Relative Abweich. (in %)	
	Mn	Cr	526	350	Mn	Cr	Mn	Cr
A1.1			0,517	1,211	0,216	0,293		
A1.2			0,517	1,212	0,216	0,293		
A1.3			0,512	1,207	0,214	0,293		
Mittel	<b>0,213</b>	<b>0,300</b>	0,515	1,210	0,216	0,293	<b>1,2</b>	<b>-2,3</b>
A2.1			0,380	1,623	0,158	0,444		
A2.2			0,380	1,623	0,158	0,444		
A2.3			0,380	1,625	0,158	0,444		
Mittel	<b>0,160</b>	<b>0,450</b>	0,380	1,624	0,158	0,444	<b>-1,1</b>	<b>-1,4</b>

#### 5.7.5 DATENBEISPIEL ZU 5.6 - MAXIMALFEHLER; KONDITIONSZAHL

Während sich der Anwender eines Kalibriermodells vor allem für den Vertrauensbereich eines Analyseergebnisses interessiert, dürfte den (akademischen) Methodenentwickler auch der Maximalfehler interessieren.

❶ Daten-ID: Modellparameter aus Messungen vom 19.12.2019; Werte siehe oben

Die relative Konditionszahl  $rcond = \|\mathbf{B}\| \|\mathbf{B}^{-1}\|$  kann am einfachsten mittels der *FROBENIUS-Normen*  $\|\mathbf{B}\|_{\text{FROB}}$  und  $\|\mathbf{B}^{-1}\|_{\text{FROB}}$  abgeschätzt werden.

Wir gehen von folgender Empfindlichkeitenmatrix aus:

$$\mathbf{B} = \begin{pmatrix} 2,410 & 0,020 \\ 1,199 & 3,200 \end{pmatrix}$$

und bilden deren Inverse:

$$\mathbf{B}^{-1} = \begin{pmatrix} 0,416 & -0,003 \\ -0,156 & 0,313 \end{pmatrix}$$

Die Werte dieser Normen ermitteln wir bei den quadratischen Matrizen aus der Summe ihrer quadrierten Elemente:

$$\|\mathbf{B}\|_{\text{FROB}} = \sqrt{\sum \sum b_{pc}^2}$$

Mit den Quadratwurzeln aus 17,486 und 0,296 erhalten wir  $rcond_{\text{FROB}} = 2,274$ .

Bei nicht-quadratischen (*P,C*)-Matrizen ermitteln wir die Werte aus der Spur von  $(\mathbf{B}^T \mathbf{B})$ , d. h. der Summe von deren Hauptdiagonalelementen, gemäß

$$\|\mathbf{B}\|_{\text{FROB}} = \sqrt{\text{spur}(\mathbf{B}^T \mathbf{B})}$$

Wir überlassen es dem Leser, die entsprechende Übung mittels der quadratischen  $\mathbf{B}$ -Matrix auszuführen.

Etwas aufwändiger ist die Berechnung der relativen Konditionszahl über die *EUKLID-oder Spektral-Norm*  $\|\mathbf{B}\|_2$ , die aus den Eigenwerten  $\eta$  nach  $\det(\eta \mathbf{I} - \mathbf{B}^T \mathbf{B}) = 0$  gebildet wird, worin  $\mathbf{I}$  die (quadratische) Einheitsmatrix ist.

$$\|\mathbf{B}\|_2 = \sqrt{\max |\eta(\mathbf{B}^T \mathbf{B})|}$$

Die Eigenwerte sind die Lösungen der charakteristischen Gleichungen für die quadratische Matrix  $\mathbf{B}^T \mathbf{B}$ . Sie sind die Nullstellen der entsprechenden charakteristischen Polynome. Mit unserem Beispiel erhalten wir aus

$$\mathbf{B}^T \mathbf{B} = \begin{pmatrix} 5,809 & 2,954 \\ 2,954 & 11,678 \end{pmatrix}$$

deren Spur übrigens 17,487 ist,

$$(\eta - b_{11}^{(T)})(\eta - b_{22}^{(T)}) - b_{12}^{(T)} b_{21}^{(T)} = 0$$

Die beiden Lösungen für  $\eta$  ergeben sich nach der quadratischen Gleichung zu:

$$\eta_{1,2} = (b_{22}^{(T)} + b_{11}^{(T)})/2 \pm [(b_{22}^{(T)} + b_{11}^{(T)})^2/4 - \det(\mathbf{B}^T \mathbf{B})]^{1/2}$$

Konkret erhalten wir somit aus

$$\eta_{1,2} = 8,74 \pm 4,164$$

die Spektralnorm

$$\|\mathbf{B}\|_2 = 3,593.$$

(Wir machen noch kurz eine Probe, denn für  $P = C = 2$  sollte gelten:

$$\|\mathbf{B}\|_2 \leq \|\mathbf{B}\|_{\text{FROB}} \leq \sqrt{2} \|\mathbf{B}\|_2.$$

Diese Probe ist mit  $3,593 \leq 4,182 \leq 5,081$  erfüllt.)

Nach MASSART et al. gilt für quadratische Matrizen, also auch für unser exakt bestimmtes Kalibriersystem  $\|\mathbf{B}\|_2 = \eta_{\text{MAX}}$  und  $\|\mathbf{B}^{-1}\|_2 = 1/\eta_{\text{MIN}}$ .

Damit ist die relative Konditionszahl der Quotient aus dem größten und kleinsten Eigenwert berechenbar. In unserem Beispiel wird  $rcond_2 = 2,818$ .

Dieser Wert ist nicht allzu unterschiedlich von  $rcond_{\text{FROB}}$  und man wird damit rechnen müssen, dass der relative Ergebnis-Fehler maximal rund das Dreifache des Mess- und Modellierfehlers betragen kann.

## 6 Mehrkomponenten-Kalibration

---

### 6.1 Modellierungsstrategie

#### 6.1.1 MULTIVARIATES STANDARD-KALIBRIERMODELL

Hierzu ist formal nicht viel Neues zu sagen, da der im vorigen Teil eingeführte Formalismus auch hier anwendbar ist: bei den formulierten Matrix- und Vektorgleichungen ist es völlig gleich, ob die Dimension  $P$  einige ausgewählte oder alle zur Verfügung stehenden Messstellen bezeichnet.

(Wir erinnern nur daran, dass Symbole für **Vektoren** und **Matrizen** immer **fett** geschrieben, und zwar mit Kleinbuchstaben für Vektoren, wobei immer Spaltenvektoren gemeint sind, und mit Großbuchstaben für Matrizen.)

Die bisher behandelten Kalibrationsmodelle sind *Modelle mit natürlichen Variablen*.

Mit natürlichen Variablen meinen wir die messtechnisch „direkt“ zugänglichen Signalgrößen. Im Unterschied dazu werden wir in den nächsten Abschnitten außerdem *latente Variablen* kennen lernen. Das sind zumeist hypothetische oder abstrakte Variablen, die ihrerseits als Linearkombinationen der natürlichen Signalgrößen oder „Original-Merkmale“ dargestellt werden können.

Erinnern wir uns zunächst daran, dass die wichtigste analytische Modellannahme war, dass ein analytisches Signal kausal durch die Konzentration einer chemischen Spezies bedingt ist. Eine weitere Annahme besteht darin, dass die mathematische Umkehrung dieser Kausalannahme einen Zugang zu der unbekanntem Konzentration ermöglicht.

Der Zweikomponenten-Fall wurde im vorigen Abschnitt besprochen.

Gehen wir nun davon aus, dass wir simultan die Konzentrationen mehrerer Analyte ( $C > 2$ ) bestimmen möchten. Rein mathematisch sind im einfachsten Fall dafür auch  $C$  Bestimmungsgleichungen erforderlich.

Die moderne Spektroskopie aber beispielsweise ermöglicht es uns in der Regel, einen „Überfluss“ von  $P > C$  Variablen im Signalraum zu erfassen.

Damit können wir das Hauptmodell der Kalibration nun in Form von linear kombinierten Beiträgen der einzelnen Analytgrößen (Konzentrationen)  $X_1, \dots, X_C$  zu jeder analytischen Signalgröße  $Y_p$  an jeweils einer Messstelle  $p$  eines bestimmten Spektralbereiches wie folgt formulieren:

$$\begin{aligned} Y_1 &= b_{11} X_1 + \dots + b_{1c} X_c + \dots + b_{1C} X_C \\ &\dots \\ Y_p &= b_{p1} X_1 + \dots + b_{pc} X_c + \dots + b_{pC} X_C \\ &\dots \\ Y_P &= b_{P1} X_1 + \dots + b_{Pc} X_c + \dots + b_{PC} X_C \end{aligned} \tag{137}$$

Der algebraische Ausdruck für eine entsprechende Messung eines *einzelnen* Gemisches aus  $C$  Analyten (Komponenten) ist identisch mit der schon früher verwendeten Formel der Kalibrierfunktion:

$$\mathbf{y} = \mathbf{B} \mathbf{x} \quad (138)$$

mit  $\mathbf{y}$   $(P,1)$ -dimensionaler Messwertevektor  
 $\mathbf{x}$   $(C,1)$ -dimensionaler Analytwertevektor (Konzentrationen)  
 $\mathbf{B}$   $(P,C)$ -dimensionale Matrix der partiellen Empfindlichkeiten

Bei der univariaten Kalibrierung ( $P = 1$ ) erhielt man, nach statistischer Absicherung der geschätzten Geradenparameter, die Analysenfunktion durch Umstellen (Invertieren) der Kalibrierfunktion.

Danach konnten die Werte  $\hat{x}$  der Analytgröße  $X$  aus jeder Signalmessung  $y$  der Messgröße  $Y$  berechnet („geschätzt“, „vorhergesagt“) werden.

Im multivariaten Fall erfolgt dies formal gemäß:

$$\hat{\mathbf{x}} = (\hat{\mathbf{B}}^T \hat{\mathbf{B}})^{-1} \hat{\mathbf{B}}^T \mathbf{y} \quad (139)$$

(Diese Möglichkeit ist auch von BROWN et al.<sup>72</sup> beschrieben worden, wird aber wegen der insgesamt zwei nötigen Matrix-Inversion nicht empfohlen.)

Das oben dargestellte Kalibriermodell kann problemlos um einen konzentrationsunabhängigen Signal-Blindwertvektor  $\mathbf{b}_0$  erweitert werden. Lediglich die Vektor- bzw. Matrixdimensionen ändern sich zu  $(P,1)$ ,  $(C+1,1)$  und  $(P,C+1)$ .

Wichtiger ist, dass man zu Bestimmung der Elemente der Empfindlichkeitenmatrix (Kalibrierkoeffizienten) nicht nur eine Mischung, sondern  $N$  Mischungen der  $C$  Analyte benötigt. Damit kommen wir, wenn wir die Mischungen formal als weitere Spalten aufschreiben, zu folgendem Gleichungssystem:

$$\mathbf{Y} = \mathbf{B} \mathbf{X} \quad (140)$$

mit  $\mathbf{Y}$   $(P,N)$ -Matrix der Signalwerte für alle Kalibrierproben  
 $\mathbf{X}$   $(C,N)$ -Matrix der Werte der Analytgröße (Konzentrationen)  
 $\mathbf{B}$   $(P,C)$ -Matrix der partiellen Empfindlichkeiten  
 $P$  Zahl der Messstellen  
 $N$  Zahl der Kalibrierproben (Gemische)  
 $C$  Zahl der verschiedenen Analytgrößen („Komponenten“)

(Bei derartigen Gleichungen sollte man immer überprüfen, ob die Dimensionen der Matrizen „zueinander passen“, ob die „Formate“ übereinstimmen, d. h. den Konventionen der Matrixalgebra folgen. Vor allem beim Vergleich mit der Literatur muss man damit rechnen, dass Spalten und Zeilen vertauscht sind.)

72 C.W. BROWN, P.F. LYNCH, R.J. OBREMSKI, D.S. LAVERY: „Matrix representations and criteria for selecting analytical wavelengths for multicomponent spectroscopic analysis“, Anal. Chem. 54(1982)1472-1479

Die Schätzung der Koeffizienten der Empfindlichkeitenmatrix erfolgt dann gemäß:

$$\hat{\mathbf{B}} = \mathbf{Y} \mathbf{X}^T (\mathbf{X} \mathbf{X}^T)^{-1} \quad (141)$$

(Das Symbol  $()^T$  bedeutet, dass die jeweilige Matrix zu transponieren ist, also die zeilen- und spaltenvertauschte Matrix verwendet wird;  $()^{-1}$  bezeichnet die Inverse einer Matrix bzw. den Vorgang der Inversion.)

Dies ist ein Ausdruck der *indirekten Kalibration* (MARTENS und NAES 1984), bei der die Empfindlichkeitenmatrix  $\mathbf{B}$  zunächst ermittelt werden muss.

Von *direkter Kalibrierung* spricht man, falls  $\mathbf{B}$  „bekannt“ ist, zum Beispiel in Form von „Spektren der reinen Komponenten“. Diese Formulierung, die man immer wieder in der Literatur findet, verschweigt, dass auch in diesem Falle  $\mathbf{B}$  erst geschätzt werden muss.

Die Vermutung liegt nahe, dass sich beide  $\mathbf{B}$ -Matrizen numerisch unterscheiden werden.

❶ Formalismus und Datenbeispiel siehe Abschnitt 6.4.1

Sowohl bei direkter als auch bei indirekter Kalibrierung unterscheidet man in der Literatur noch zwischen dem natürlichen Fall (zufällige Kalibrierung) und dem *kontrollierten* Fall (P. J. BROWN 1982), je nachdem, ob die Kalibrierproben eine Zufallsstichprobe der gesamten (möglichen) Probenpopulation darstellen oder ob sie systematisch und unter Kontrolle ausgewählt wurden. Für die Vorhersage von Konzentrationen impliziert dies spezielle statistische Eigenschaften (NAES und MARTENS 1984). Eine Übersicht der multivariaten linearen Kalibriermethoden auf statistischer Basis gibt NAES (1985).

Wir haben schon häufig auf den großen (numerischen) Vorteil hingewiesen, der sich im Kalibrierschritt durch ein („kontrolliertes“) Versuchsdesign nach einem (*orthogonalen*) statistischen Versuchsplan ergibt. Dann ist nämlich die zur Schätzung von  $\mathbf{B}$  zu invertierende Matrix eine Einheitsmatrix.

Ein wichtiger Spezialfall liegt natürlich dann vor, wenn  $\mathbf{B}$  eine quadratische (C,C)-Matrix ist ( $P = C$ ). Ihre Koeffizienten entsprechen dann den partiellen Empfindlichkeiten und  $\mathbf{B}^{-1}$  direkt den Koeffizienten der Analysenfunktionen.

## 6.1.2 MULTIVARIATES INVERSES KALIBRIERMODELL

Nicht nur aus numerischen Gründen kann es sinnvoll sein, ein „inverses“ Kalibriermodell folgenden Typs anzunehmen:

$$\mathbf{X} = \mathbf{B} \mathbf{Y} \quad (142)$$

Dies ist der *P-Matrix-Ansatz* (vgl. BROWN und OBREMSKI 1984).

Die Schätzung der Werte in  $\mathbf{B}$  kann bei Vorliegen genügend vieler Messungen wiederum mittels multivariater multipler linearer Regressionsanalyse (MMR, MLR) erfolgen.

$$\hat{\mathbf{B}} = \mathbf{X} \mathbf{Y}^T (\mathbf{Y} \mathbf{Y}^T)^{-1} \quad (143)$$

Diese Formulierung wird von den genannten Autoren aus numerischen Gründen bevorzugt, da nur eine Matrixinversion ausgeführt werden muss.

Es wurde früher aber daran kritisiert, dass man damit im Sinne der Statistik nur dann verlässliche Schätzwerte erhält, wenn angenommen werden kann, dass die Messfehler der Signale kleiner als diejenigen der Konzentrationen sind.

Heute, zum Beispiel speziell in der NIR-Spektroskopie, erweist sich gerade das inverse Kalibriermodell als mathematisch eher begründet als das klassische Modell. Der Grund liegt darin, dass tatsächlich die Analytgrößen bzw. ihre Realisierungen in Form der Kalibrierproben weitaus stärker fehlerbehaftet sein können als die Signalmessungen.

❶ Datenbeispiel siehe Abschnitt 6.4.1

## 6.2 Problemvermeidungsmöglichkeiten

### 6.2.1 NICHTKONSTANZ DER KALIBRIERKOEFFIZIENTEN

Kann keine Konstanz der Kalibrierkoeffizienten über den gesamten Kalibrierbereich erwartet werden, zum Beispiel durch chemische „Matrixeffekte“ (etwa in der Spurenanalytik), so gewinnt die generalisierte Standardadditionsmethode (GSAM) an Bedeutung. Es könnte auch ein physikalischer „Matrixeffekt“, etwa infolge zeitlicher Signaldrift, sich problematisch auswirken.

Speziell auf den Fall des nichtkonstanten Untergrundes beim Übergang von Kalibrierproben zu Analysenproben geht die verallgemeinerte Untergrundsubtraktion (GBS, generalized background subtraction) ein. Im univariaten Fall konnte dem Problem meist dadurch begegnet werden, dass man Messungen der Blindwerte beiderseits der eigentlichen Messstelle vornahm. Beim multivariaten Fall ergibt sich die Korrektur quasi von selbst. Die simultane mehrdimensionale Mehrkomponenten-Analytik enthält diese Korrektur explizit nicht. (Man vergleiche dazu das in Abschnitt 6.2.5 Gesagte.)

### 6.2.2 NUMERISCHE PROBLEME

Die Messwertematrix  $\mathbf{Y}$  (Rohdaten) kann gelegentlich fehlerbehaftet sein, insbesondere dann, wenn der Rauschanteil der Spektren die Effekte der kausalen Einflussfaktoren  $\mathbf{X}$  quasi nivelliert, d. h. die reinen Komponenteneffekte im Rauschen untergehen. Dadurch evtl. hervorgerufene *Multikollinearitäten* können zu numerischen Problemen führen, nämlich bei mathematischen Verfahren, die Matrixinversionen erfordern. Im schlimmsten Fall können diese Multikollinearitäten zu Instabilitäten der Berechnungsalgorithmen führen und die Zuverlässigkeit der Ergebnisse beeinträchtigen.

Im „X-Block“ der experimentellen Variablen kann die Gefahr durch Multikollinearitäten durch handwerklich fehlerhafte Wahl der Kalibrierproben hervorgerufen werden. Zum Beispiel führen einfache Verdünnungsreihen eines einzigen Mehrkomponenten-Gemi-

sches automatisch zu linearen Abhängigkeiten der Konzentrationen und damit zu Problemen durch singuläre Matrizen.

Schon im univariaten Fall, bei der Anwendung einer multiplen Regression, würde die schlechte Kondition der Korrelationsmatrix der  $X$ -Variablen auffallen. Man könnte dann die Regressionsparameter nicht mehr stabil mittels OLS schätzen, sondern müsste beispielsweise auf die „Ridge Regression“ ausweichen. (Sie erniedrigt künstlich die Korrelationskoeffizienten.)

Eventuell würde auch eine schrittweise MLR ausreichend sein.

Diese Probleme lassen sich von Anfang an vermeiden, indem man die Konzentrationsverhältnisse nach statistischen Versuchsplänen wählt. Darauf haben wir immer wieder hingewiesen, denn für uns ist dies der „Königsweg“ (siehe Abschnitt 6.3.1).

Auch die Spektren schlecht geplanter Versuchsansätze können dann linear abhängig sein und singuläre Matrizen bewirken.

Im „Y-Block“ der experimentellen Variablen kann man Multikollinearitäten durch wirksame *Datenvorbehandlung* beseitigen. Die danach angewandten Kalibriermodelle beziehen sich dann meist nicht mehr auf den Raum der Signalvariablen, sondern auf einen Raum latenter Variablen. Diese Variablen können im Datenvorbereitungsschritt beispielsweise durch eine mathematische Hauptkomponentenzerlegung erzeugt werden. (Siehe dazu auch den folgenden Abschnitt 6.2.3.)

Damit sichert man, dass die ehemals linear abhängigen Zeilen bzw. Spalten einer Rohdatenmatrix in orthogonale Zeilen bzw. Spalten umgewandelt werden. Man führt dann mit diesen statt der üblichen multivariaten linearen Regression (MLR) eine Hauptkomponentenregression (PCR; engl.: principal components regression) durch.

Mathematisch ist das Problem dadurch charakterisiert, dass eine Matrix  $\mathbf{Z}$  nicht von vollem Rang ist, wobei manchmal zwischen Spalten- und Zeilenrang unterschieden wird. Dann werden Berechnungen von  $(\mathbf{Z}^T \mathbf{Z})^{-1}$  numerisch instabil.

Bei den Kalibrationsmethoden tritt diese Inversionsaufgabe bei allen drei Matrixtypen auf: in der „traditionellen“ Kalibration mit  $\mathbf{X}$  und  $\mathbf{B}$ , in der „pragmatischen inversen“ Kalibration sowie bei explorativen multivariaten Anwendungen mit  $\mathbf{Y}$  (vgl. die beiden folgenden Abschnitte).

Beispielsweise bei AHRENS und LÄUTER<sup>73</sup> findet man den Hinweis, wie man den Rang  $C$  einer  $(N,C)$ -Matrix  $\mathbf{Z}$  einfach bestimmen kann. Wenn deren  $C$  Spalten linear unabhängig sind, dann ist die folgende Hilfsmatrix (*Projektionsmatrix*)  $\mathbf{H}$

$$\mathbf{H} = \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \quad (144)$$

idempotent mit dem Rang  $C$ . Dieser ergibt sich rechnerisch schnell aus der Spur von  $\mathbf{H}$ , ist also gleich der Summe der Hauptdiagonalelemente:

---

73 H. AHRENS, J. LÄUTER: „Mehrdimensionale Varianzanalyse. Hypothesenprüfung. Dimensionserniedrigung. Diskrimination“, Akademie-Verlag, Berlin, 1981, S. 7, 13 und 17

$$C = \text{rang}(\mathbf{H}) = \text{spur}(\mathbf{H}) = \sum_{c=1}^C h_{ii} \quad (145)$$

In gleicher Weise kann man den Rang  $P$  einer  $(N,P)$ -Matrix  $\mathbf{Z}$  bestimmen.

❶ Datenbeispiel siehe Abschnitt 6.4.3

### 6.2.3 RAUSCHANTEIL BESEITIGEN (LATENTE VARIABLEN)

Eine Möglichkeit, den Rauschanteil der Spektren zu isolieren, d.h. die Redundanz in der Datenmatrix zu beseitigen und mit einer (transformierten) Datenmatrix zu arbeiten, deren Elemente nur durch das Wirken „echter“ Faktoren zustande gekommen sind, bietet sich durch Zerlegung der Rohdatenmatrix  $\mathbf{Y}$ :

$$\mathbf{Y} = \mathbf{A}_Y \mathbf{F}_Y + \mathbf{E}_Y \quad (146)$$

mit  $\mathbf{Y}$  (Signal-)Datenmatrix  
 $\mathbf{F}$  latente Variablen (abstrakte Faktoren)  
 $\mathbf{A}$  Anteile dieser Variablen in den einzelnen Spektren (Objekten).  
 $\mathbf{E}$  Matrix der Restfehler („Rauschanteil“).

Diese Zerlegung ist möglich mittels mathematischer Hauptkomponentenanalyse (PCA), auch Hauptkomponentenzerlegung oder Hauptachsentransformation. Sie hat eine enge Beziehung zur Faktoranalyse (FA), denn bei beiden Methoden erfolgt zunächst die Berechnung der Eigenvektoren der Kovarianzmatrix  $\mathbf{S} = \mathbf{Y}'\mathbf{Y}'$ , wobei  $\mathbf{Y}'$  die zentrierten (mittelwertkorrigierten) Werte von  $\mathbf{Y}$  enthält.

Durch Rangermittlung, das heißt die Entscheidung über die signifikanten Eigenvektoren und damit über eben die latenten Faktoren, die auf chemische Komponenten (reine Stoffe mit ihren Spektren) zurückzuführen sind, kann dann der Rauschanteil separiert und mit  $\mathbf{E}$  vereinigt werden.

Zur Schätzung der Zahl der Faktoren in der PCA bzw. FA stehen *Kreuzvalidierungstechniken* oder empirische Funktionen zur Verfügung. Wenn nur wenige Hauptkomponenten ermittelt werden sollen, also keine komplette Hauptkomponentenzerlegung erforderlich ist, wurde das *NIPALS*-Verfahren (engl.: non linear iterative partial least squares) von WOLD als rechentechnisch schnell und einfach empfohlen.

Lässt sich  $\mathbf{Y}$  durch eine a-priori-Information bereits in Teilblöcke zerlegen, kann zur weiteren Suche nach abstrakten Faktoren die *Diskriminanzanalyse* (DA) angewendet werden. Wir werden auf diese statistisch fundierte Vorgehensweise in Abschnitt 6.3.2 eingehen.

Um auch Fehler in der Konzentrationsmatrix  $\mathbf{X}$  zu berücksichtigen, kann diese analog zu  $\mathbf{Y}$  zerlegt werden:

$$\mathbf{X} = \mathbf{A}_X \mathbf{F}_X + \mathbf{E}_X \quad (147)$$

Diesen Weg beschreitet das *PLS-Verfahren* von WOLD<sup>74</sup> ursprünglich für zwei Datenblöcke  $\mathbf{Y}$  und  $\mathbf{X}$ . (PLS ist hier zu lesen als: projection to latent structures.)

Der Nutzen derartiger Zerlegungstechniken wird augenfällig, wenn man bedenkt, dass die moderne Analysetechnik zunehmend Datenmatrizen höherer Ordnung produziert. Mittels generalisierter Rang-Annihilations-Faktoranalyse (RAFA), im Wesentlichen eine Folge von PCA, hat man sogar Lösungen für dreidimensionale Matrizen publiziert.<sup>75</sup>

Sowohl bei der Behandlung von Multikollinearitäten als auch bei der Verbesserung des Signal-Rausch-Verhältnisses in der Multivariaten Mehrkomponenten-Analytik wird als Kriterium die *Konditionszahl* empfohlen (vgl. Abschnitt 5.7.5). Ein direkter Vergleich von Konditionszahlen der Originaldaten-Matrizen mit denen der Matrizen, die im Raume der latenten Variablen definiert sind, ist jedoch nicht sehr aussagekräftig, zum Beispiel sind beide Konditionszahlen bei vollständiger Faktorzerlegung gleich, haben aber unterschiedliche Bedeutung.

## 6.2.4 VISUELLE AUSREIßERERKENNUNG

Im Teil „Einkomponenten-Kalibration“ wurde bereits darauf hingewiesen, dass Ausreißer sowohl die Modellierung als auch das analytische Ergebnis verfälschen können. In der dort meist univariat betriebenen Kalibration genügen in der Regel einfache graphische Methoden, um Ausreißer zu erkennen.

❶ Datenbeispiel siehe Abschnitt 6.4.4

Bei der multivariaten Kalibration, die in einem höherdimensionalen Signalraum stattfindet, lassen sich ebenfalls graphische Methoden nutzen.

Die einfachsten dabei sind solche, die meist auf die Analysenproben und Ergebnisvariablen bezogen sind. Dazu zählen zum Beispiel Sterne-Diagramme, Polar-Diagramme und die CHERNOFF-Gesichter.

Um Analyseobjekte im Raum ihrer transformierten Original-Messgrößen abzubilden, eignen sich *Projektionsmethoden*. Zu diesen wollen wir auch die *Clusteranalyse* mit ihren Ergebnissen in Form von *Dendrogrammen* zählen.

Vor allem aber erweisen sich die Projektionen im Anschluss an eine mathematische *Hauptkomponentenanalyse* (*Hauptachsentransformation*) oder eine *Faktoranalyse* als vorteilhaft.

❶ Datenbeispiel siehe Abschnitt 6.4.4

---

74 S. WOLD, P. GELADI, K. ESBENSEN, J. ÖHMANN: „Multi-Way Principal Components and PLS-Analysis“ J. Chemom. 1(1987)41-56

75 Beispiele für dreidimensionale Datenmatrizen sind zum Beispiel {Proben x GC-Läufe x Massenspektren} oder {Proben x LC-Läufe x UV-Spektrum} oder {Proben x Anregungsfrequenz x Emissionsspektrum (Fluoreszenz)}. Beispiele für Matrizen der Ordnung 4 ergeben sich bei Kombinationen wie zum Beispiel {Proben x Chromatogramm x Säule x Spektrum} oder {Proben x 2-D-Chromatogramm x UV-Spektrum}.

## 6.2.5 „WECHSELWIRKUNGEN“ UND „INTERFERENZEN“

In diesem Abschnitt geht es um die *simultane Berücksichtigung* von Konzentrations- und Signalwechselwirkungen, die meist etwas lapidar und unspezifisch als Interferenzen bezeichnet werden.

Es wurde bereits darauf hingewiesen, dass bei der direkten Schätzung der Matrix der Koeffizienten der Analysenfunktionen praktisch jede Analysenfunktion unabhängig von den anderen ermittelt wird und gleiches für die Koeffizienten der Kalibrierfunktionen gilt.

Mit Modellen aus der Statistischen Versuchsplanung können dabei gewisse spektrale bzw. Analyt-Interferenzen jeweils eindimensional (sequentiell, ungekoppelt) berücksichtigt werden. Siehe dazu Abschnitt 6.3.1, wo die Optimierung der Kalibrierprobenwahl im Vordergrund steht.

Effekte von Wechselwirkungen (im Sinne von Multikollinearitäten) im Bereich der Einflussgrößen können, wie bereits angemerkt, auch durch den Übergang zu latenten Variablen umgangen werden.

Es gibt aber auch die Möglichkeit einer „Systemschätzung“ von Koeffizienten mittels der dreistufigen Least-Squares-Methode. So erlaubt das Modell der Strukturgleichungen (SEQ) explizit die Verwendung hochkorrelierter Einflussgrößen. Es liegt nahe, die entsprechenden effizienten Schätzmethode auch in der Mehrkomponenten-Kalibration und -Analyse anzuwenden.

❶ Datenbeispiel siehe Abschnitt 6.4.3

## 6.3 Aufwandsoptimierung

### 6.3.1 KALIBRIERPROBEN OPTIMAL HERSTELLEN

Hierfür eignet sich besonders die *Statistische Versuchsplanung*, wenn man von einer begrenzten Anzahl von Analytgrößen ausgehen kann.

Die Attraktivität der statistischen Versuchsplanung besteht darin, die Parameter von Zielfunktionen (im statistischen Sinne) optimal zu schätzen, wobei sie nur ein Minimum an Experimentalaufwand erfordert.

Wenn diese Parameter die gleichartigen, auf wechselnde Konzentration mehrerer Analyte bezogenen partiellen Empfindlichkeiten einer Kalibrierfunktion sind, kann man auf diese Weise „unliebsamen“ Konzentrationswechselwirkungen begegnen.

In der üblichen Darstellung der Versuchsplanung führen die  $M$  Kombinationen (Mischungen, Versuche) der diskreten Stufen der  $C$  Einflussgrößen  $X$  (Analyte) zu einer Matrix  $\mathbf{X}_{VP}$ . In jedem dieser  $M$  Versuche erhält man den Antwortvektor *einer* Zielgröße  $Y_z$  (mit  $z = 1, \dots, Z$ ).  $Z$  ist die Anzahl der Antwort- oder Zielgrößen, die der Anzahl  $P$  unserer früheren Modelle entspricht.

$$\mathbf{y}_z = \mathbf{X}_{VP} \mathbf{b}_z \quad (148)$$

mit  $\mathbf{y}_z$  ( $M,1$ )-Signal-Vektor einer quasi „kontrollierten Stichprobe“  
 $\mathbf{X}_{VP}$  ( $M,C$ )-Matrix der Mischungen  
 $\mathbf{b}_z$  ( $C,1$ )-Vektor.

Mit  $M \geq C$  Versuchen lassen sich die Parameter der Beziehung  $Y_z = f(X_1, \dots, X_C)$  in bekannter Weise ermitteln:

$$\mathbf{b}_z = (\mathbf{X}_{VP}^T \mathbf{X}_{VP})^{-1} \mathbf{X}_{VP}^T \mathbf{y}_z \quad (149)$$

Das heißt, jeder Koeffizient in  $\mathbf{b}_z$  kann so unabhängig von den anderen Koeffizienten geschätzt werden. Im statistischen Sinne sind diese Koeffizienten unkorreliert und bei einer normalverteilten Zielgröße auch unabhängig.

Auf die Beziehung der Versuchsplanmatrix  $\mathbf{X}_{VP}$  bzw. des zu invertierenden Produkts zur „Informationsmatrix“ eines Versuchsplans gehen wir im Datenbeispiel in Abschnitt 6.4.5 ein.

Schließlich können mit derselben Versuchsplan-Matrix, d. h. in praxi, aus denselben  $M$  Versuchen sogar weitere, eigentlich beliebig viele Zielgrößen modelliert werden, sofern deren Antwortvektoren  $\mathbf{y}_z$  bekannt sind oder gemessen werden können.

Bei ausreichend vielen Kalibrierproben kann man mittels der statistischen Versuchsplanung für jede Zielgröße außerdem gleichzeitig *nichtlineare* Modellansätze für Analyt-Wechselwirkungen (in der Konzentrationsdomäne) formulieren.

❶ Datenbeispiel siehe Abschnitt 6.4.4

### 6.3.2 MESSSTELLENBEWERTUNG UND -AUSWAHL

Im nächsten Abschnitt werden wir eine *statistisch fundierte* Möglichkeit der Messstellenauswahl besprechen. Aber nicht jeder Interessierte will vielleicht gleich mit dieser statistisch anspruchsvollen Methode beginnen. Wir empfehlen deshalb die vergleichsweise einfache, „rein mathematische“ Methode der Hauptkomponentenzerlegung der Signalmatrix  $\mathbf{Y}$ , die bereits in Abschnitt 6.2.3 erwähnt wurde.

Dabei geht man von folgender Zerlegung der Signalmatrix  $\mathbf{Y}$  aus:

$$\mathbf{Y} = \mathbf{A} \mathbf{F} + \mathbf{E} \quad (150)$$

mit  $\mathbf{Y}$  ( $N,P$ )-Datenmatrix  
 $\mathbf{A}$  ( $N,S$ )-Matrix der Anteile („Gewichte“) der latenten Variablen in den einzelnen Objekten (Spektren) (engl.: factor loadings)  
 $\mathbf{S}$  Anzahl der Faktoren (latenten Variablen)  
 $\mathbf{F}$  ( $S,N$ )-Matrix der Faktorwerte eines jeden Objekt (Spektrum) (engl.: scores)  
 $\mathbf{E}$  ( $N,P$ )-Matrix des „Rekonstruktionsfehlers“

Der Rest- oder Rekonstruktionsfehler entsteht hier dadurch, dass wir eine Dimensionsreduktion für die Darstellung der Objekte erzielen wollen:  $S < P$ .

Anstelle einer Darstellung, beispielsweise unserer Kalibrierproben, im hochdimensionalen Raum von Dutzenden oder Hunderten von Wellenlängen arbeiten wir im niedrigerdimensionalen Raum von nur  $S$  „latenten“ Größen. Wir können sogar hoffen, dass  $S \ll P$  ist.

❶ vgl. das Datenbeispiel in Abschnitt 6.4.5

Welche Wellenlängen in diesen Größen eine besondere Rolle spielen, werden wir daraus entnehmen können, wie hoch die „Ladungen“ der Original-Messgrößen (Wellenlängen) ausfallen.

Bei künftigen Messungen kann man auf „weniger wichtige“ Messstellen dann verzichten. Darauf beruht die Idee der Messstellenauswahl mittels mathematischer Hauptkomponentenanalyse.

Die erwähnte Zerlegung von  $\mathbf{Y}$  wird durch „Hauptachsentransformation“ der Kovarianzmatrix  $\mathbf{S}$  oder der Korrelationsmatrix  $\mathbf{R}$  der Rohdaten ausgeführt.

Dazu sucht man erneut die nichttrivialen Lösungen  $\eta$  des Eigenwertproblems, das wir bereits im Abschnitt 5.6 kennen gelernt haben.

$$| \mathbf{R} - \eta \mathbf{I} | = 0 \tag{151}$$

mit  $\mathbf{R}$   $(P,P)$ -dimensionale Korrelationsmatrix  
 $\mathbf{I}$   $(P,P)$ -dimensionale Einheitsmatrix  
 $\eta$  Eigenwert

Zu jeder Lösung des Eigenwertproblems

$$\mathbf{R} \mathbf{e}_p = \eta_p \mathbf{e}_p \tag{152}$$

gehört ein Lösungspaar („Eigenpaar“) aus Eigenwert und zugehörigem Eigenvektor:

$\eta_p$  Eigenwert  
 $\mathbf{e}_p$  Eigenvektor  
 $p = 1, 2, \dots, P$ .

Wichtig für uns ist, dass folgendes gilt:

$$\sum_{p=1}^P \eta_p = P \tag{153}$$

Da jeder Eigenwert zugleich proportional zu einem Varianzanteil ist, der in  $\mathbf{R}$  steckt, kann mittels einer graphischen Darstellung (*Screeplot*) der numerisch geordneten Eigenwerte entschieden werden, welchen Varianzanteil man mit  $S < P$  Faktoren modellieren (rekonstruieren) will.

❶ vgl. das Datenbeispiel in Abschnitt 6.4.6

### 6.3.3 SYSTEMISCHE SIGNAL/RAUSCH-OPTIMIERUNG

Schon 1988 wurde vorgeschlagen, die lineare Diskriminanzanalyse (LDA) zu verwenden, um eine „systemische Signal-Rausch-Optimierung“ von Kalibriermodellen zu erreichen.<sup>76</sup> In der quantitativen Analytik ist man bemüht, das Nutzsignal im Vergleich zum Rauschen zu maximieren. Eine ähnliche Optimierungsaufgabe hat die Diskriminanzanalyse als Mustererkennungsmethode: Es wird die optimale Darstellung von Datenklassen mit Hilfe der Diskriminanzfunktionen (DF) gesucht, die das Verhältnis von *faktorbedingter zu zufälliger Varianz* berücksichtigen.

In der Analytik ist der Faktor die Variation der Einflussgröße „Konzentration“ einer Komponente bzw. des „Konzentrationsverhältnisses“ mehrerer Komponenten. Durch diese Variation erhofft man eine signifikante (kausale) Wirkung dergestalt, dass sich die Signalgröße(n), zum Beispiel in Form aufgenommener Spektren, deutlich (signifikant) ändern. Die Versuche oder Spektren, die zu jeder Faktorstufe gehören, bilden jeweils eine Klasse von Messwerten. Der Unterschied zwischen diesen  $K$  Klassen sollte also faktorbedingt sein.

Die faktorbedingte Varianz wird durch eine Matrix der *Zwischen-Klassen-„Streuungen“*  $\mathbf{B}$  erfasst.  $\mathbf{B}$  (von engl.: between) darf hier nicht mit der Empfindlichkeitenmatrix  $\mathbf{B}$  verwechselt werden.

Die Variation der Werte in den einzelnen, durch wiederholte Messung derselben Mischung entstehenden Klassen sollte in der Regel zufallsbedingt sein. Sie wird durch eine Matrix der *Inner-Klassen-„Streuung“*  $\mathbf{W}$  (von engl.: within) dargestellt.

$$\mathbf{B} = \sum_{k=1}^K N_k (\bar{\mathbf{y}}_k - \bar{\mathbf{y}})(\bar{\mathbf{y}}_k - \bar{\mathbf{y}})^T$$

$$\mathbf{W} = \sum_{k=1}^K \sum_{n=1}^{N_k} (\mathbf{y}_{kn} - \bar{\mathbf{y}}_k)(\mathbf{y}_{kn} - \bar{\mathbf{y}}_k)^T$$
(154)

mit  $\mathbf{B}$  ( $P,P$ )-Matrix der Zwischen-Klassen-Abweichungen  
 $\mathbf{W}$  ( $P,P$ )-Matrix der Inner-Klassen-Abweichungen  
 $K$  Anzahl der Kalibrierklassen (Faktorstufenkombinationen)  
 $N_k$  Anzahl der Versuche in Klasse  $k$ ,  $k = 1, \dots, K$   
 $\bar{\mathbf{y}}$  Gesamtmittelwertvektor  
 $\bar{\mathbf{y}}_k$  Klassenmittelwertvektor

Damit wird die signalbezogene Datenvariation im (varianzanalytischen) Sinne des Effektes der Komponentenvariation explizit in  $\mathbf{B}$  berücksichtigt, während die messfehlerbedingte Variation in  $\mathbf{W}$  erfasst wird.

Somit sind die „Grundgrößen“ der Optimierungsaufgabe definiert.

<sup>76</sup> H. W. ZWANZIGER: „Mustererkennung und Mehrkomponentenanalytik mit multivariaten chemometrischen Methoden“, Habilitationsschrift, Universität Leipzig, Leipzig, 1988

Die LDA ermittelt nun diejenigen Linearkombinationen der natürlichen Signalgrößen, mit denen eine simultane Trennung aller Klassen im statistischen Sinne gleich gut (optimal) möglich ist. Diese Linearkombinationen werden dann die Unterschiede zwischen den Klassen im Vergleich zu den Unterschieden innerhalb der Klassen maximieren.

Das der LDA zugrunde liegende Extremalproblem läuft mathematisch auf die Lösung eines allgemeinen Eigenwertproblems hinaus, das wir so formulieren können:

$$| \mathbf{B} - \eta \mathbf{W} | = 0 \quad (155)$$

Das spezielle Eigenwertproblem würde man so formulieren:

$$| \mathbf{B} \mathbf{W}^{-1} - \eta \mathbf{I} | = 0 \quad (156)$$

Die Auflösung des Eigenwertproblems führt zu  $T$  Eigenwerten  $\eta_1, \dots, \eta_T$ . (Wir verwenden, wie schon in Abschnitt 5.6, den griechischen Buchstaben  $\eta$  statt des in der Literatur üblichen Buchstabens  $\lambda$ , um eine Verwechslung mit dem Wellenlängensymbol  $\lambda$  zu vermeiden.)

Zu jedem Eigenwert gehört jeweils ein Eigenvektor  $\mathbf{e}$ , dessen Werte die Koeffizienten der *nichtelementaren*<sup>77</sup> linearen Diskriminanzfunktionen (LDF) sind. Diese Werte sind somit Wichtungsfaktoren der  $P$  Originalgrößen.

Analog zur Hauptkomponentenzerlegung im vorigen Abschnitt ist konkret für die ersten beiden Lösungen des Eigenwertproblems zu schreiben:

$$\begin{aligned} (\mathbf{B}\mathbf{W}^{-1}) \cdot \mathbf{e}_1 &= \eta_1 \cdot \mathbf{e}_1 \\ (\mathbf{B}\mathbf{W}^{-1}) \cdot \mathbf{e}_2 &= \eta_2 \cdot \mathbf{e}_2 \end{aligned} \quad (157)$$

Der erste Eigenwert entspricht dem Anteil der Varianz in  $\mathbf{B}\mathbf{W}^{-1}$ , die durch den ersten Eigenvektor  $\mathbf{e}_1$  „extrahiert“ wird, der zweite Eigenwert beschreibt die Varianz, die der zweite Eigenvektor extrahiert usw. Die ungewichteten Koeffizienten der Eigenvektoren entsprechen jeweils den „Gewichten“ der Originalmessgrößen.

Wichtig bzw. vorteilhaft ist, dass die Anzahl der von Null verschiedenen Eigenwerte und damit die Zahl  $T$  der Diskriminanzfunktionen durch die Klassenzahl  $K$  (Konzentrationsstufen) und die Anzahl  $P$  der Messgrößen gegeben ist:

$$T = \min (K-1, P) \quad (158)$$

Da die Klassenzahl im allgemeinen viel kleiner ist als  $P$ , lässt sich damit schließlich auch jede Messprobe statt im Raum ihrer  $P$  Originalmessgrößen in einem wesentlich niedrigerdimensionalen Raum darstellen. Damit eröffnen sich einerseits erfreulich anschauliche graphische Darstellungsmöglichkeiten der Proben.

---

<sup>77</sup> Der Begriff wird für Diskriminanzfunktionen verwendet, zu deren Ermittlung ein Eigenwertproblem zu lösen ist.

Andererseits ermöglichen es methodentypische Algorithmen und Trennmaße, die Zahl der für die optimale Klassentrennung nötigen Original-Messgrößen zu reduzieren. Für bestimmte multivariate Analysenaufgaben ist damit unmittelbar ein verminderter Messaufwand verbunden.

Eine häufig genutzte Testgröße für die (Trenn-)Wirksamkeit solcher Teilmengen der Original-Größen ist „WILKS' Lambda“

$$\lambda_{WILKS} = \frac{|W|}{|T|} \quad (159)$$

mit  $|W|$  Determinante der Inner-Klassen-Streuung-Matrix  
 $|T|$  Determinante der Gesamtstreuung-Matrix

Damit berechnet man ein „multiplikatives Inkrement“, das partielle  $\lambda_{PART}$ , in Form des Quotienten aus dem Wert der Testgröße, nachdem und bevor man eine Variable der Trennmenge hinzugefügt hat. Aus dieser Testgröße wird dann ein empirischer  $F$ -Wert berechnet und gegen einen kritischen  $F$ -Wert (Zufallshöchstwert) getestet. Auf Details gehen wir hier nicht ein.<sup>78</sup>

❶ Datenbeispiel zur Aufwandsoptimierung siehe auch Abschnitt 6.4.6

Die Idee der „systemischen Signal/Rausch-Modellierung“ besteht nun darin, die Eigenvektoren von  $BW^{-1}$ , die in gewisser Weise „Systemeigenschaften“ abbilden, zu verwenden, um „verrauschte“ Originaldaten geeignet zu transformieren und danach deren Werte den üblichen Kalibrationsprozeduren zu unterwerfen.

Die Verfahrensschritte der MLR-Anwendung auf Originaldaten bzw. auf transformierte Daten kann man für die Kalibrier- und die Anwendungsphase wie folgt schematisieren. Zur Symbolik sei erinnert bzw. ergänzt:

<b>X</b>	(N,C)-Matrix der C Einflussgrößen (Konzentrationen)
<b>Y</b>	(N,P)-Matrix der Messwerte
<b>Y'</b>	(N,P)-Matrix mit leerwertkorrigierten Werten
<b>U</b>	(P,T)-Matrix von „unstandardisierten“ Eigenvektoren
<b>B</b> <sub>&lt;Index&gt;</sub>	(C,P)-Matrix der Empfindlichkeiten des Kalibriersystems
<b>A</b>	(N <sub>ANW</sub> ,P)-„Anwendungsmatrix“; ergibt sich aus <b>B</b> <sub>&lt;Index&gt;</sub>

Algorithmische *Rechenverfahren* (symbolisch:  $\Rightarrow$ VERFAHREN  $\Rightarrow$ ) sind von *Matrixoperationen* (symbolisch:  $\rightarrow$ ) zu unterscheiden.<sup>79</sup>

Verfahrensschritte mit Originaldaten:

Kalibrieren:

$X, Y_{ORIG} \Rightarrow MLR \rightarrow Y'_{ORIG}$  und  $B_{ORIG}$

$B_{ORIG} \rightarrow A_{ORIG}$

78 Genauer findet man in den Handbüchern der entsprechenden Software, die derartige Berechnungen ausführt, zum Beispiel in STATISTICA-Handbuch (1994) S. 3095

79 Der Pfeil  $\rightarrow$  ist also immer zu lesen als: „angewandt auf“ bzw. „führt zu“.

Anwenden:

$$Y_{\text{ANW}} := Y'_{\text{ORIG}} \mid \mathbf{A}_{\text{ORIG}} \rightarrow Y_{\text{ANW}} \rightarrow X_{\text{PRED}} \text{ (berechne SEC)}$$

Verfahrensschritte mit transformierten Daten:

Kalibrieren:

$$K, Y_{\text{ORIG}} \xrightarrow{\text{LDA}} \mathbf{U}_{\text{TC}}$$

$$\mathbf{U}_{\text{TC}} \rightarrow Y_{\text{ORIG}} \rightarrow Y_{\text{TC}}$$

$$X, Y_{\text{TC}} \xrightarrow{\text{MLR}} Y'_{\text{TC}} \text{ und } \mathbf{B}_{\text{TC}}$$

$$\mathbf{B}_{\text{TC}} \rightarrow \mathbf{A}_{\text{TC}}$$

Anwenden:

$$Y_{\text{ANW}} := Y'_{\text{TC}} \mid \mathbf{A}_{\text{TC}} \rightarrow Y_{\text{ANW}} \rightarrow X_{\text{PRED}} \text{ (berechne SEC)}$$

$$Y_{\text{ANW}} := Y_{\text{ORIG}} \mid \mathbf{A}_{\text{TC}} \rightarrow Y_{\text{ANW}} \rightarrow X_{\text{PRED}} \text{ (berechne SEP)}$$

❶ Datenbeispiel siehe auch Abschnitt 6.4.7



Modell für	$b_1$ (Mn)	$b_2$ (Cr)	Leerwert
$\lambda_1$	0,935	2,341	0,031
$\lambda_2$	1,238	3,100	0,042
$\lambda_3$	2,374	0,016	0,000
$\lambda_4$	2,283	0,015	0,000

Völlig vergleichbar können beide Ergebnisse schon deswegen nicht sein, weil diese Matrix nach der Vermessung von *Komponentengemischen* ermittelt wurde, was auch zu den Leerwerten führt.

#### 6.4.2 DATENBEISPIEL ZU 6.1.2 - INVERSE KALIBRATION

Numerisch gesehen, ist die inverse Kalibration vorzuziehen. Allerdings wird sie in der Praxis weniger benutzt.

- ❶ Daten-ID: 19.12.2019; M-Proben, Design: 2<sup>2</sup>-VP mit Zentralpunkt, Mittelwerte aus Dreifachmessungen bei zwei Wellenlängen

Mit den bekannten Daten (siehe Abschnitt 1.7.2) ergeben sich folgende Ergebnisse. Die Determinante des inversen Kalibriersystems ist mit 0,1 zwar recht klein, aber unbedenklich, da wir die Empfindlichkeitenmatrix zum Zwecke der Anwendung nicht invertieren müssen.

Die anderen Ergebnisse sprechen für sich.

Modellparameter (Koeffizienten)	Leerwert		Empfindlichkeit ( $\lambda_1$ )		Empfindlichkeit ( $\lambda_2$ )	
$Y(\text{Mn}) = f(\lambda_1, \lambda_2)$	$b_{10}$	0,006	$b_{11}$	0,416	$b_{12}$	-0,003
$Y(\text{Cr}) = f(\lambda_1, \lambda_2)$	$b_{20}$	-0,009	$b_{21}$	-0,155	$b_{22}$	0,313

Fehler der Modellparameter	Leerwert		Empfindlichkeit ( $\lambda_1$ )		Empfindlichkeit ( $\lambda_2$ )		$B$	$S_{\text{RES}}$	$SEC^*$	$SEP^*$
Modell $Y(\text{Mn})$	$s(b_{10})$	0,000	$s(b_{11})$	0,000	$s(b_{12})$	0,000	1,000	0,000	0,772	0,002
Modell $Y(\text{Cr})$	$s(b_{20})$	0,006	$s(b_{21})$	0,007	$s(b_{22})$	0,004	0,998	0,007	11,14	0,007

\* Die Berechnung des „Kalibrierfehlers“  $SEC$  wäre obsolet, da keine „Vorhersage“ von Signalwerten angestrebt wird. Dass der Vorhersagefehler der Konzentrationsgrößen so gut ausfällt, überrascht auch nicht, da sie Gegenstand der Kleinste-Quadrate-Schätzung der Koeffizienten waren.

#### 6.4.3 DATENBEISPIEL ZU 6.2.2 - LINEAR ABHÄNGIGE KALIBRIERDATEN

Linear abhängige Kalibrierdaten sind möglichst zu vermeiden, da sie numerische Probleme bereiten können. Aus unserer Sicht entstehen linear abhängige Daten im X-Bereich durch schlechte Versuchsplanung. Im Y-Bereich können, wenn schon nicht als Folge der schlechten Konditionierung der X-Matrix, so evtl. durch „rauschbedingte Signalnivellierung“ lineare Abhängigkeiten entstehen.

- ❶ Daten-ID: 02.02.2020; M4-Proben; bewusstes Fehl-Design: Dreipunktkalibration, 5 Wellenlängen, Dreifachmessungen; Konzentrationen in mmol l<sup>-1</sup>

Wir machen hier bewusst den „handwerklichen“ Fehler und Verdünnen die Probe M4 zweimal (v1 und v2), womit wir folgende Signaldaten erhalten. Am besten wird die lineare Abhängigkeit sichtbar, wenn man die Mittel der Signalwerte mit zwei Dezimalstellen angibt. Die eigentlichen Rechnungen führen wir aber mit voller Zahlenlänge bzw. den neun Einzelmesswerten aus.

Probe	Konzentration		Absorbanz bei den Wellenlängen $\lambda$ (in nm)				
	Mn	Cr	290	312	350	526	544
M4 mean	0,32	0,450	1,40	1,16	1,83	0,76	0,72
M4v1 mean	0,16	0,225	0,75	0,59	0,92	0,38	0,36
M4v2 mean	0,08	0,1125	0,27	0,29	0,46	0,19	0,18

### (1) Vorahnung

Dass derartige Verdünnungsreihen zu numerischen Problemen mindestens bei Rechenverfahren der „traditionellen“ Kalibration führen müssen, zeigt das folgende allgemeine Beispiel einer Dreipunktkalibration. Wir stellen durch Verdünnen einer Stammlösung mit den Startkonzentrationen „4a“ und „4b“ zweier Analyte drei Kalibrierproben her, mittels derer die Empfindlichkeitenmatrix bei verschiedenen Wellenlängen berechnet werden soll.

Probe	Konzentration (X)		Absorbanz (Y) bei				
	$X_1$	$X_2$	$\lambda_1$	...	$\lambda_p$	...	$\lambda_p$
1	4a	4b	...	...	...	...	...
2	2a	2b	...	...	...	...	...
3	a	b	...	...	...	...	...

Aus dem Kalibriermodell  $\mathbf{Y} = \mathbf{X}\mathbf{B}$  erhält man durch Umformen die Matrix  $\mathbf{B}$ :  $\mathbf{B} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$ . Für die zu invertierende Matrix  $(\mathbf{X}^T \mathbf{X})$  ergibt sich

$$(\mathbf{X}^T \mathbf{X}) = 21 \begin{pmatrix} a^2 & ab \\ ab & b^2 \end{pmatrix}$$

Da deren Determinante gleich Null ist, kann die Rechnung nicht fortgeführt werden. Kalibrieransätze dieser Art müssen also scheitern.

### (2) Bestätigung

STATISTICA meldet konsequent für jede Wellenlänge eine „schlecht konditionierte Matrix“.

### (3) Ausweg für $\mathbf{X}$ ?

Unser „handwerklicher Fehler“ ist offenbar so gravierend, dass auch die *Ridge Regression* uns nicht weiter bringt. Zwar funktioniert der Algorithmus klaglos, findet aber für keine Wellenlänge eine Kalibriergleichung mit signifikanten Koeffizienten.

Wir haben drei Durchläufe mit je unterschiedlichen Werten (0,1 bis 0,3) für die Stellkonstante probiert, die die Korrelationsmatrix verändert. In keinem Fall konnten signifikante Parameter geschätzt werden.

Die  $\mathbf{B}$ -Matrix, auf ein paar Validier-Testproben angewendet, liefert folglich völlig unakzeptable Konzentrationswerte. Wir verzichten hier auf die Wiedergabe numerischer Werte.

#### (4) Prüfen von $\mathbf{Y}$

Wie nicht anders zu erwarten, ist auch die  $\mathbf{Y}$ -Matrix nicht von vollem Rang. Da schon die Determinante von  $(\mathbf{Y}^T \mathbf{Y})$  Null ist, führen wir den nächsten Schritt zur Rangbestimmung, die Inversion, gar nicht erst aus.

#### 6.4.4 DATENBEISPIEL ZU 6.2.4 - VISUELLE AUSREIßERERKENNUNG ( $\lambda$ -PAARE, CA, PCA)

Auf die Bedeutung der Ausreißererkenung wurde schon mehrfach hingewiesen. Im  $X$ -Raum sollten Ausreißer bei statistischer (faktorieller) Versuchsplanung ausgeschlossen sein. Im ein- und zweidimensionalen  $Y$ -Raum sind viele grafische Methoden zur Ausreißererkenung von Kalibrierproben bekannt. Man kann sich beispielsweise der Mühe unterziehen, die Proben in jeder Ebene zu betrachten, die jeweils von zwei der insgesamt  $P$  Signalvariablen (Wellenlängen) aufgespannt wird. Das wären aber  $P(P-1)/2$  Bilder, von denen wir zwei im Teil (1) vergleichen.

Im drei- und vor allem im höherdimensionalen Signalraum müssen wir auf andere Methoden zurückgreifen, die uns in nur einem Bild die „volle Information“ vermitteln. Als voraussetzungsfree Methoden werden wir im folgenden Teil (2) die Clusteranalyse (CA) und in Teil (3) die Hauptkomponentenanalyse (PCA) besprechen. Beide Methoden erlauben zugleich, Beziehungen zwischen den Messgrößen zu visualisieren. Da wir es immer mit Gruppen (gleiches Konzentrationsverhältnis zweier Analyte) zu tun haben, wären auch Methoden geeignet, die dieses Vorwissen einbeziehen können. Die dafür besonders infrage kommende Diskriminanzanalyse behalten wir uns für andere Anwendungen vor (siehe Abschnitt 6.3.6).

##### ❶ Daten-ID:

Teil (1): 21.01.2020; Mischungsproben, Design:  $3^2$ -VP, 5 Wellenlängen; Einfachmessungen; daraus simulierte Mehrfachmessungen; Konzentrationen in  $\text{mmol l}^{-1}$

Teil (2) und (3): 21.01.2020; Mischungsproben, Design:  $3^2$ -VP, 5 Wellenlängen, Einfachmessungen; Konzentrationen in  $\text{mmol l}^{-1}$

Probe	Konzentration		Absorbanz bei Wellenlänge $\lambda$ (in nm)				
	Mn	Cr	290	312	350	526	544
M11	0,08	0,15	0,465	0,342	0,590	0,194	0,184
M21	0,20	0,15	0,588	0,543	0,731	0,482	0,461
M31	0,32	0,15	0,708	0,744	0,871	0,772	0,739
M12	0,08	0,30	0,842	0,554	1,072	0,197	0,185
M22	0,20	0,30	0,966	0,756	1,216	0,489	0,465
M32	0,32	0,30	1,088	0,958	1,357	0,775	0,740
M13	0,08	0,45	1,211	0,770	1,554	0,200	0,186
M23	0,20	0,45	1,336	0,973	1,699	0,486	0,461
M33	0,32	0,45	1,454	1,170	1,846	0,776	0,740

#### (1) Zweidimensionale grafische Darstellungen

Beim systematischen Durchmustern aller bivariaten Streudiagramme könnte man bereits erkennen, bei welchen Wellenlängen-Kombinationen sich die Struktur der Versuchspunkte am deutlichsten zeigt. Von den 10 möglichen Wellenlängenkombinationen vergleichen wir hier die Wellenlängenkombination 290|312 nm (zwei Cr-Peaks) mit der Kombination 350|544 nm (die Peak-Maxima im Cr- und im Mn-Spektrum).

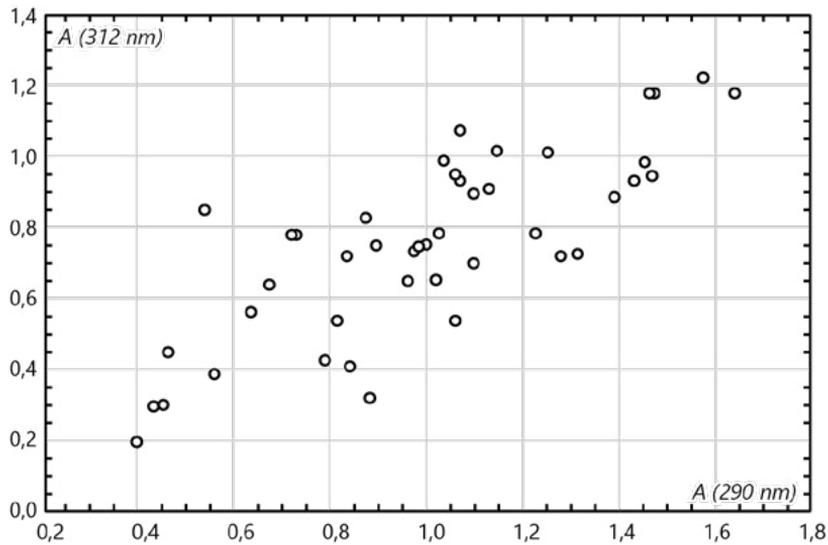


Abb. 14: Abbildung der Versuchspunkte (suboptimales Wellenlängenpaar)

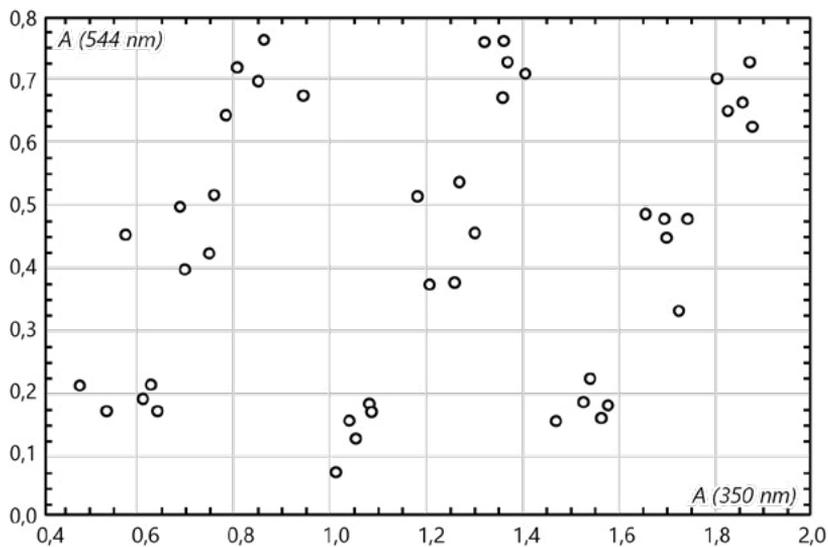


Abb. 15: Abbildung der Versuchspunkte (optimales Wellenlängenpaar)

## (2) Anwendung der Clusteranalyse (CA)

Eine Clustering von Rohdaten setzt voraus, dass die Variablen gleiche „Skaleneinheiten“ besitzen. Anderenfalls sind die Daten vor der Clustering zum Beispiel einer *Autoskalierung* bzw. *Standardisierung* zu unterwerfen. Bei dieser speziellen Skalierung werden die Werte (variablenweise) zunächst durch Abzug des Mittelwerts zentriert und dann durch die Standardabweichung dividiert.

In den einschlägigen Programmen kann man zwischen den Clustering-Typen „k-means“ und „agglomerativ“ wählen. Im ersten Falle wird die Zahl der erwünschten (zu findenden) Cluster (Gruppierungen, Klassen) vorgegeben. Wir kennen zwar unsere Klassenzahl,

nämlich die 9 Versuchspunkte des 3<sup>2</sup>-Versuchsplans, aber wir wollen „unvoreingenommen“, also „explorativ“ vorgehen und entscheiden uns für die zweite Variante. Außerdem maximiert das erste Vorgehen die Zwischen-Cluster-Abstände, was im Falle einfach besetzter Klassen, zumindest in der Startphase, schiefgehen könnte.

Im nächsten Schritt können wir festlegen, ob wir Fälle oder Variablen klassifizieren wollen. In der Regel interessieren uns die *Fälle*.

Zwischen den Fällen (Objekten) werden *Abstände* berechnet, auf die jeder Clusterungsalgorithmus nun zugreift. (Vorteilhaft für andere Forschungsdisziplinen ist, dass man statt der Einzeldaten auch von einer beliebig gewonnenen Distanzmatrix ausgehen kann.)

Der einfachste räumliche Abstand, den wir durch unsere eigene „Raumerfahrung“ kennen, ist der dreidimensionale EUKLID-Abstand. Er ist praktisch das räumliche Pendant zum PYTHAGORAS-Abstand zweier Punkte in der (zweidimensionalen) Ebene, den wir aus der Schulzeit kennen.

Im  $P$ -dimensionalen Raum unserer Messgrößen definieren wir zunächst den EUKLIDischen Abstand  $d_{ij}$  zweier Proben  $i$  und  $j$  gemäß der ersten Gleichung in der folgenden Formel.

Mit diesen Werten erhält man eine  $(N,N)$ -Distanzmatrix aller Kalibrierproben.

Nun werden die beiden ähnlichsten Objekte fusioniert, zum Beispiel mögen die Objekte  $i$  und  $j$  das neue Objekt  $k$  bilden, das wir Cluster  $k$  nennen. Dessen Klassenabstand zu jedem verbleibenden Objekt  $l$  (oder Cluster) muss jetzt neu berechnet werden. Damit ergibt sich eine veränderte Distanzmatrix.

Die folgende zweite Formel verwendet dazu eine Reihe von Faktoren ( $\alpha$ ,  $\beta$  und  $\gamma$ ), durch deren Werte sich typische hierarchische Clusterverfahren unterscheiden.

$$d_{ij} = \sqrt{\sum_{p=1}^P (y_{ip} - y_{jp})^2} \tag{160}$$

$$k := i + j$$

$$d_{kl} = \alpha_i d_{il} + \alpha_j d_{jl} + \beta d_{ij} + \gamma |d_{il} - d_{jl}|$$

Diese Abstandsmaße sollten in gewisser Weise zum Verfahren passen. So passt zum EUKLID-Abstand das „älteste und zugleich einfachste Verfahren“<sup>81</sup>, das Verfahren „single linkage“. Es ist übrigens vor allem dann interessant, wenn es um die „Erkennung verzweigter oder gekrümmter, linienförmiger oder kreisförmiger Gebilde“ geht.

Will man ähnliche *Variablen* zusammenfassen, geht man vom Distanzmaß „1 – PEARSONSs  $r$ “ aus und probiert zuerst die Strategie „single linkage“.

Für „Standardanwendungen“ der Clusteranalyse zur Mustererkennung haben wir häufig mit gutem Erfolg das WARDSche Verfahren benutzt, zu dem quadrierte EUKLIDische

---

81 D. STEINHAUSEN, K. LANGER: „Clusteranalyse. Einführung in Methoden und Verfahren der automatischen Klassifikation. Mit zahlreichen Algorithmen, FORTRAN-Programmen, Anwendungsbeispielen und einer Kurzdarstellung der multivariaten statistischen Verfahren“ (einschließlich eines Abschnitts „Grundbegriffe aus der Linearen Algebra“), de Gruyter Verlag, Berlin, New York, 1977

Distanzen am besten passen, nach anderen Erfahrungen noch besser aber deren halbierte Werte.<sup>82</sup>

Die Distanzwerte lassen sich in Form von *Dendrogrammen* grafisch darstellen. An den drei Hauptzweigen des Baumdiagramms sehen wir deutlich, dass die Struktur der Daten zuerst durch Chrom beeinflusst wird. Erst die weitere Verästelung zeigt den Einfluss der Mangan-Konzentration.

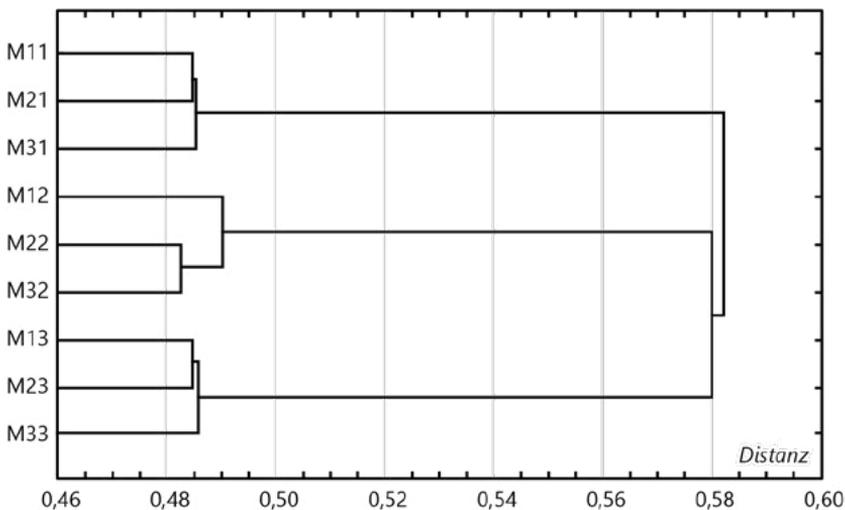


Abb. 16: Dendrogramm der 9 Versuchspunkte

Das Dendrogramm der Wellenlängen zeigt (bis zur Distanz 0,1) eindeutig die drei charakteristischen Chrom-Wellenlängen in einem Cluster und die beiden Mangan-Wellenlängen im zweiten Cluster.

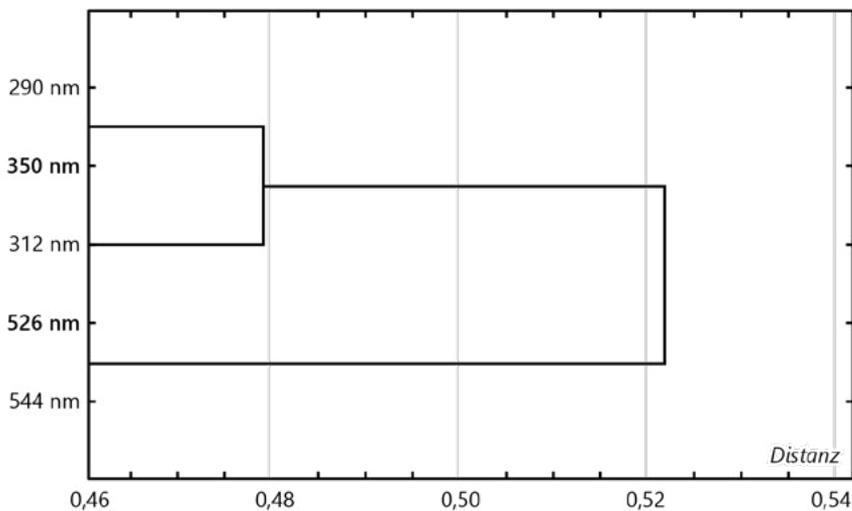


Abb. 17: Dendrogramm der Wellenlängen

82 H.-J. MUCHA: „Minireport CLUSTER“, Akademie der Wissenschaften der DDR, Karl-Weierstraß-Institut für Mathematik, Berlin, 1986, S. 20

### 3. Anwendung der Hauptkomponentenanalyse (PCA)

Unsere Rohdatenmatrix wird von „gleichartigen“ Variablen gebildet, wie oben am Beginn des Absatzes über die Clusteranalyse konstatiert. Daher könnten wir von der Kovarianzmatrix  $S$  ausgehen. Wir gehen hier aber den üblichen Weg: Aus den standardisierten  $Y$ -Variablen wird die Korrelationsmatrix  $R$  berechnet, die wir der Hauptkomponentenzerlegung unterwerfen werden.

Diese Matrix, deren Werte in Tabellenform folgen, zeigt auch, welche Wellenlängenwerte signifikant korrelieren. (Diese Werte sind fett hervorgehoben.)

	290 nm	312 nm	350 nm	526 nm	544 nm
290 nm	1,000	<b>0,905</b>	<b>1,000</b>	0,318	0,312
312 nm	<b>0,905</b>	1,000	<b>0,893</b>	<b>0,690</b>	<b>0,686</b>
350 nm	<b>1,000</b>	<b>0,893</b>	1,000	0,291	0,285
526 nm	0,318	<b>0,690</b>	0,291	1,000	<b>1,000</b>
544 nm	0,312	<b>0,686</b>	0,285	<b>1,000</b>	1,000

Auf die numerischen Werte der Zerlegung von  $R$  gehen wir im Detail nicht ein.

Wir sehen schon an den Eigenwerten (EW) und dem zugehörigen charakteristischen „Screepplot“, dass die ersten beiden (mathematischen) Hauptkomponenten den Hauptanteil der Datenvarianz erfassen.

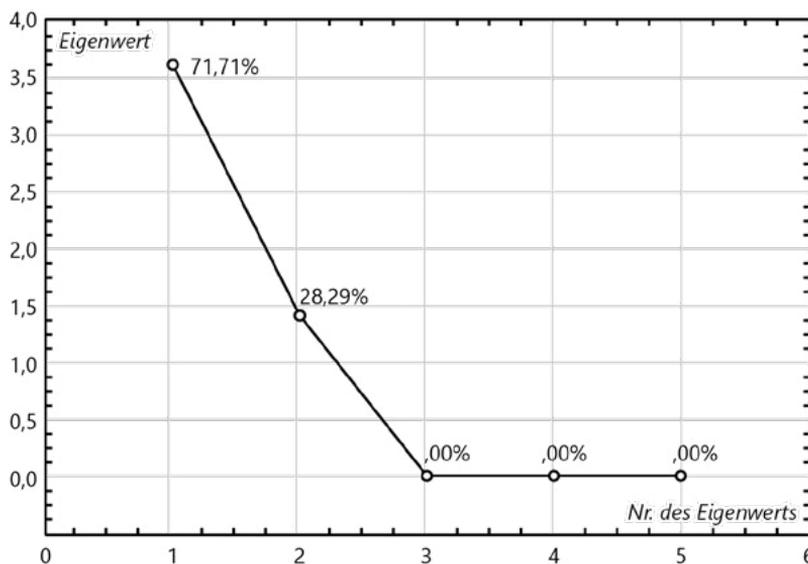


Abb. 18: Screepplot zur Hauptkomponentenanalyse

Die numerischen Verhältnisse verdeutlicht folgende Tabelle, die neben den Eigenwerten (Einzel- und kumulativer Wert) deren Beitrag zur Datenvarianz auflistet (% Gesamt- und kumulativer Wert):

Nr.	Eigenwert	Kumul. EW	% Gesamt	Kumul.
1	3,585	3,585	71,707	71,707
2	1,415	5,000	28,291	99,998
3	0,000	5,000	0,001	99,999
4	0,000	5,000	0,001	100,000
5	0,000	5,000	0,000	100,000

Deshalb fällt auch die Projektion unserer Versuchspunkte aus dem fünfdimensionalen Raum der Wellenlängen mittels ihrer Faktorkoordinaten auf die Faktorebene der beiden Hauptfaktoren bilderbuchartig aus und zeigt die Güte des Versuchsplan-Ansatzes und natürlich keinerlei „Ausreißer“.

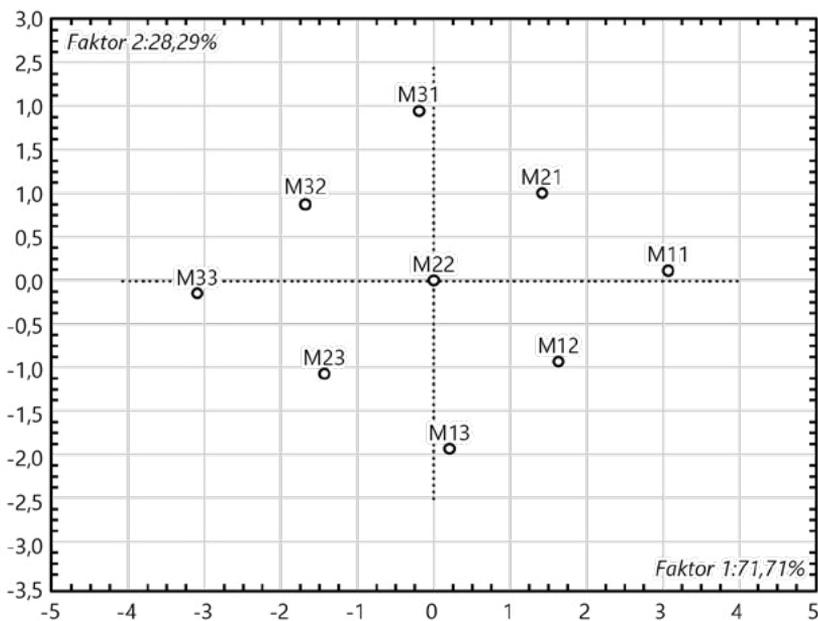


Abb. 19: Projektion der Kalibrierversuche mittels Hauptfaktoren

Ganz ähnlich lässt sich aus den Faktorladungen der Wellenlängen deren grafische Darstellung erhalten, also eine Projektion des fünfdimensionalen Wellenlängen-Raums auf die Ebene der beiden Hauptkomponenten-Achsen. Wir erkennen in dieser Darstellung auch sehr schön die Ähnlichkeiten (Cluster) der beiden Wellenlängen-Mengen für Chrom und Mangan sowie die Sonderstellung des Peakminimums von Chrom.

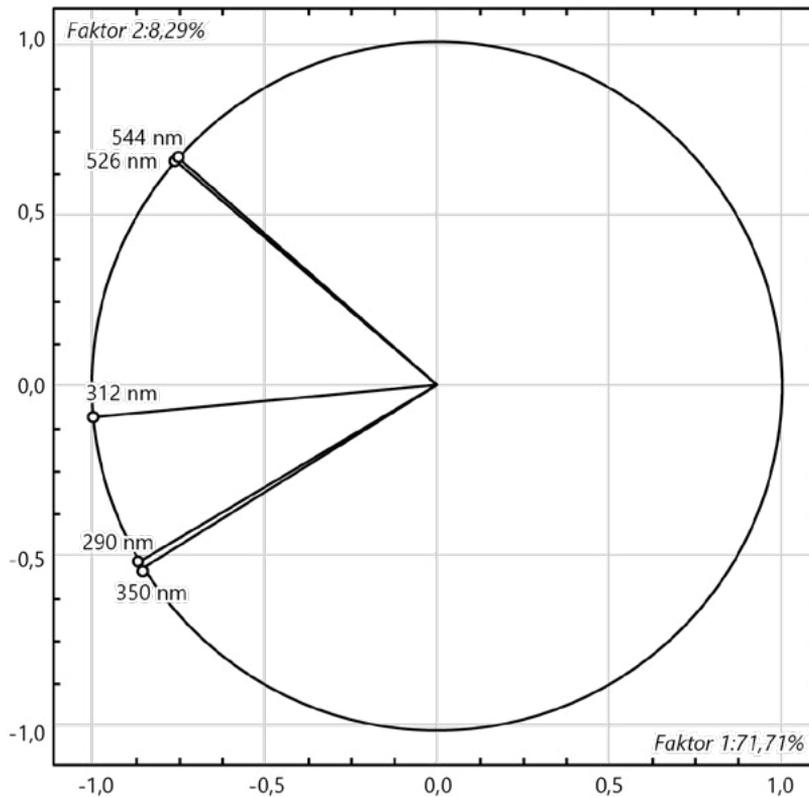


Abb. 20: Projektion der Wellenlängen mittels Hauptfaktoren

Im Rahmen der hier diskutierten Ausreißererkenntnis ist das hier ohne Bedeutung, im Zusammenhang mit der Messstellenauswahl kommen wir aber darauf zurück.

#### 6.4.5 DATENBEISPIEL ZU 6.3.1 - OPTIMALE KALIBRIERPROBEN (VERSUCHSPLANUNG)

Um experimentellen Aufwand zu sparen, haben wir mehrfach vorgeschlagen, nach einem statistischen Versuchsplan (VP) vorzugehen. Am einfachsten wählt man dafür die höchsten und niedrigsten Analytkonzentrationen  $X$  als Konzentrationsstufen aus und kombiniert diese nach einem zweifaktoriellen, zweistufigen Faktorplan. In Teil (1) verwenden wir seine Daten analog einer „Zweipunktkalibration“. Erweitert um eine Kombination der beiden mittleren Analytwerte (Zentralpunkt) wäre der Versuchsplan auch im Sinne einer Dreipunktkalibration nutzbar. Dafür am besten geeignet ist ein zweifaktorieller, dreistufiger Versuchsplan  $3^2$ , den wir in Teil (2) verwenden.

In der Schreibweise mit geeignet transformierten Konzentrationswerten sind diese Versuchspläne für beliebige ähnliche Sachverhalte verwendbar.

Zur grafischen Darstellung der Messwerte eignen sich prinzipiell die in Abschnitt 6.4.4 besprochenen Methoden.

❶ Daten-ID:

Teil (1): 25.10.2019; M-Proben, Design:  $2^2$ -VP ohne Zentralwert, Zweipunktkalibration, Mittelwerte aus Dreifachmessungen; Konzentrationen in  $\text{mmol l}^{-1}$

Teil (2): 21.01.2020; Mischungsproben; Design:  $3^2$ -VP, 5 Wellenlängen, Einfachmessungen; Konzentrationen in  $\text{mmol l}^{-1}$

(1) Zweifaktorieller, zweistufiger Versuchsplan  $2^2$

Probe	Konzentration		transform. Konzentration	
	Mn	Cr	Mn	Cr
M1	0,08	0,15	-1	-1
M2	0,08	0,45	-1	1
M3	0,32	0,15	1	-1
M4	0,32	0,45	1	1

Mit den Werten für „half range“,  $hr = (x_{\text{MAX}} - x_{\text{MIN}})/2$ , bei Mn: 0,12 und bei Cr: 0,15, sowie für „central value“,  $cv = (x_{\text{MAX}} + x_{\text{MIN}})/2$ , bei Mn: 0,2 und bei Cr: 0,3, erhält man die transformierten Konzentrationen für jeden Ausgangswert gemäß:

$$x_{\text{TRANS}} = (x - cv)/hr.$$

Nur zur Daten-Identifikation seien auch die Absorbanzen bei folgenden Wellenlängen  $\lambda$  (in nm) aufgeführt:

	526	350
M1	0,182	0,608
M2	0,185	1,505
M3	0,743	0,861
M4	0,749	1,786

Man bezeichnet diese Einser-Darstellung auch als den *Kern des Versuchsplans* und die zugehörige Matrix mit  $\mathbf{Z}$ . Mit den zu diesen vier Versuchen gehörenden Signalen (einer oder mehrerer Signalsgrößen  $Y$ ) könnte man die vier Parameter einer Gleichung vom Typ

$$Y = f(X_1, X_2) \text{ bzw.}$$

$Y = f(Z_1, Z_2) = b_0 + b_1 Z_1 + b_2 Z_2 + b_{12} Z_1 Z_2$  schätzen. Dazu erweitern wir die  $\mathbf{Z}$ -Matrix zur Auswertematrix  $\mathbf{A}$ .

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

(Man beachte, dass die Versuche damit immer noch eindeutig codiert sind.)

Die erste Spalte wird später zur Berechnung von  $b_0$ , die letzte zur Berechnung von  $b_{12}$  benötigt. Der letztgenannte Term erlaubt die Abschätzung von möglichen „Konzentrationswechselwirkungen“. Für eine statistisch fundierte Bewertung benötigt man allerdings Mehrfachmessungen.

Mittels des Matrixprodukts ( $\mathbf{A}^T \mathbf{A}$ ) definiert man die *Informationsmatrix* des Versuchsplans. Wenn diese eine Diagonalmatrix bildet, liegt ein orthogonaler Versuchsplan vor. Dies trifft, wie man selbst nachrechnen kann, auf unser Beispiel zu.

Es sei noch angemerkt, dass im Falle einer zufälligen Wahl ausreichender Versuche eine passende Visualisierungsmethode (wie im vorigen Beispiel CA, PCA oder gar FA) helfen kann, die „Ausgewogenheit“ bzw. Ausreißerfreiheit der Kalibrierproben zu beurteilen. (Siehe das „Bilderbuchergebniss“ bei  $3^2$ -VP mit PCA bzw. FA im vorigen Datenbeispiel.)

## (2) Zweifaktorieller, dreistufiger Versuchsplan $3^2$

Erweitern wir das Konzentrationsschema der Dreipunkt-Kalibration für einen Analyten wie folgt um einen zweiten Analyten, könnte unser Versuchsplan folgende Proben enthalten. (Wiederholung der Daten aus Abschnitt 6.4.4.)

Probe	Konzentration		Absorbanz bei Wellenlänge $\lambda$ (in nm)				
	Mn	Cr	290	312	350	526	544
M11	0,08	0,15	0,465	0,342	0,590	0,194	0,184
M21	0,20	0,15	0,588	0,543	0,731	0,482	0,461
M31	0,32	0,15	0,708	0,744	0,871	0,772	0,739
M12	0,08	0,30	0,842	0,554	1,072	0,197	0,185
M22	0,20	0,30	0,966	0,756	1,216	0,489	0,465
M32	0,32	0,30	1,088	0,958	1,357	0,775	0,740
M13	0,08	0,45	1,211	0,770	1,554	0,200	0,186
M23	0,20	0,45	1,336	0,973	1,699	0,486	0,461
M33	0,32	0,45	1,454	1,170	1,846	0,776	0,740

Der Leser versuche selbst, das entsprechende Z-Schema aufzustellen.

### 6.4.6 DATENBEISPIEL ZU 6.3.2 - MESSSTELLENAUSWAHL (FA, LDA)

Für diese Aufgabe eignet sich (indirekt) die Hauptkomponenten- bzw. Faktoranalyse (FA), die wir bereits als explorative Methode in Abschnitt 6.4.4 kennenlernten. Wir nutzen ihre Ergebnisse in Teil (1). Zu einer statistisch begründbaren (direkten) Messstellenauswahl empfehlen wir die (lineare) Diskriminanzanalyse (LDA), die das Vorwissen um die Gruppen gleicher Konzentrationsverhältnisse beider Analyte explizit (und konfirmativ) einbezieht. Ihre Ergebnisse sind Gegenstand von Teil (2).

❶ Daten-ID: Teil (1): 21.01.2020; Mischungsproben, Design:  $3^2$ -VP, 5 Wellenlängen, Einfachmessungen; Konzentrationen in  $\text{mmol l}^{-1}$

Probe	Konzentration		Absorbanz bei Wellenlänge $\lambda$ (in nm)				
	Mn	Cr	290	312	350	526	544
M11	0,08	0,15	0,465	0,342	0,590	0,194	0,184
M21	0,20	0,15	0,588	0,543	0,731	0,482	0,461
M31	0,32	0,15	0,708	0,744	0,871	0,772	0,739
M12	0,08	0,30	0,842	0,554	1,072	0,197	0,185
M22	0,20	0,30	0,966	0,756	1,216	0,489	0,465
M32	0,32	0,30	1,088	0,958	1,357	0,775	0,740
M13	0,08	0,45	1,211	0,770	1,554	0,200	0,186
M23	0,20	0,45	1,336	0,973	1,699	0,486	0,461
M33	0,32	0,45	1,454	1,170	1,846	0,776	0,740

❶ Teil (2): Daten-ID: 21.01.2020; Mischungsproben, Design:  $3^2$ -VP, 5 Wellenlängen; Einfachmessungen; daraus simulierte Mehrfachmessungen; Konzentrationen in  $\text{mmol l}^{-1}$

#### (1) Faktoranalyse (FA)

Diese Methode zur (explorativen) Datenanalyse startet gewöhnlich von der Korrelationsmatrix  $\mathbf{R}$  der Rohdaten. Die Faktoranalyse kann einfach zur Datenreduktion durch „Extraktion der Hauptkomponenten“ verwendet werden. Sie kann auch in Form der „ausgefeilterten“ Hauptfaktoranalyse Strukturen in den Beziehungen der Messgrößen finden.

(Manche Programme bieten bei der FA als Extraktionsmethoden die „Hauptachsenmethode“ oder die „Kommunalitätenmethode“ an. Unter der *Kommunalität* einer Messgröße versteht man denjenigen Anteil ihrer Varianz, der durch die extrahierten Hauptfaktoren erklärt wird. Der verbleibende Anteil wäre ihre spezifische Varianz.)

Wir entscheiden uns nach den Erkenntnissen aus der PCA-Anwendung in 6.4.4 gezielt für die Berechnung (Extraktion) von nur zwei Faktoren (Hauptachsen) und wählen in unserer Software die „Extraktionsmethode: Hauptkomponenten“.

Die erhaltenen Eigenwerte sind die gleichen wie im Abschnitt 6.4.4 Teil (3).

Aus der grafischen Darstellung der Faktorladungen erkennen wir bereits die „Entbehrlichkeit“ der Wellenlänge 312 nm, deren Position sich klar von den beiden Wellenlängengruppen mit Analytbezug Cr bzw. Mn unterscheidet.

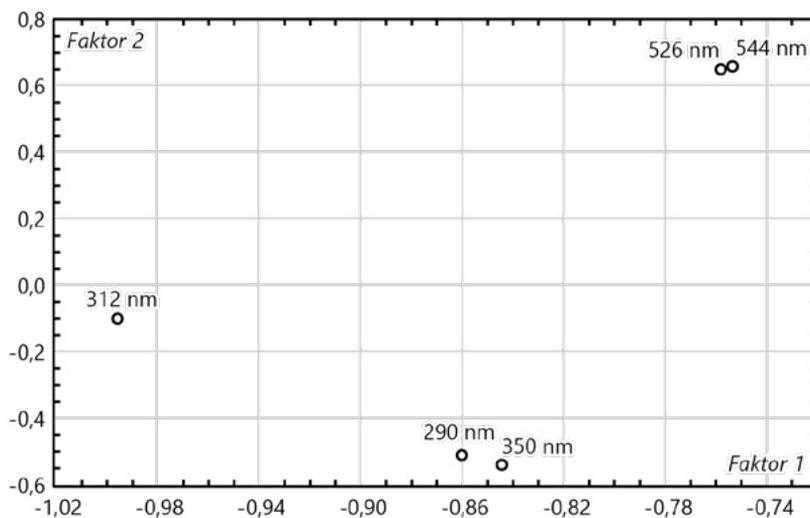


Abb. 21: Faktorladungen (unrotiert) der Faktoranalyse

Man könnte nun noch eine Varimax-Rotation der Faktoren anschließen, die selbst bei noch mehr Wellenlängen ein klares optisches Bild des Charakters der ermittelten Faktoren liefern würde.

Die numerischen Werte der rotierten Faktorladungen zeigen allerdings eindeutig, dass der erste Faktor  $F_1$  vor allem durch die drei Cr-Wellenlängen 290, 312 und 350 nm „aufgeladen“ wird, während der zweite Faktor  $F_2$  durch die beiden Mn-Wellenlängen 526 und 544 nm bestimmt wird.

$\lambda_p$	$F_1$	$F_2$
290	0,990	0,141
312	0,836	0,548
350	0,994	0,113
526	0,181	0,983
544	0,175	0,984

Man könnte also aus diesen latenten Wellenlängenfaktoren jeweils die Wellenlänge bzw. Wellenlängen auswählen, die für einen der Faktoren repräsentativ ist. Die Auswahlmengen sind für  $F_1$  {290, 312, 350 nm} und für  $F_2$  {526, 544}. Auf diese Weise könnte diese Methode prinzipiell für die Aufgabe der *Messstellenauswahl* aus einem wesentlich größeren Angebot von Wellenlängen zu brauchen sein.

Die Faktoranalyse, der zu Beginn dann natürlich keine Faktoranzahl vorgeschrieben werden sollte, erwiese sich damit als echte explorative Methode der Datenanalyse, konkret zum Entdecken von Ähnlichkeiten (Beziehungen) zwischen Variablen.

Dabei wäre explizit kein „Vorwissen“ über eine Klassenstruktur der Daten erforderlich, wie es bei der Methode im nächsten Abschnitt der Fall ist. Aber warum sollten wir nicht unser Zusatzwissen zu den Daten verwenden, um vielleicht sogar schärfere Aussagen treffen zu können?

## (2) Lineare Diskriminanzanalyse (LDA)

Während wir oben (PCA, FA) nur einer Vermutung über „passende“ Wellenlängen nachgehen konnten, bietet die LDA die Möglichkeit einer statistisch begründeten Messstellen-Auswahl. Aus der Erfahrung wissen wir, dass dabei manchmal überraschende Messstellen-Kombinationen auftreten können, mit denen eine optimale Trennung der vorgegebenen Datenklassen erfolgen kann.

### (2.1) Simulierte Daten

Wir verwenden die gleichen Daten wie oben, müssen allerdings für jeden Versuchspunkt weitere Messungen simulieren. Nur bei Mehrfachbesetzungen jeder Klasse (Faktorkombination) kann die hier empfohlene Diskriminanzanalyse jeweils eine „Inner-Klassen-Streumatrix“ berechnen und verwenden.

Um das Verfahren noch weiter zu strapazieren, haben wir den Datensatz in STATISTICA auf 45 Fälle (5 Wiederholungen der X-Werte) aufgeweitet und die Ausgangswerte für Y zunehmend mit normalverteilten Zufallszahlen aus dem Intervall von *Null* bis *Max* verfälscht.

Mit diesem Absorbanz-„Rauschen“ simulieren wir in gewisser Weise eine Art Messfehler. *Max* durfte folgende Werte annehmen: 0,005 („kaum“ Rauschen), 0,05 („mäßiges“ Rauschen), 0,1 („starkes“ Rauschen) und 0,5 („extremes“ Rauschen).

### (2.2) Aufbau- oder Abbauverfahren?

Man kann darüber philosophieren, ob man sich der Ausführung der Diskriminanzanalyse für das Aufbau- oder das Abbauverfahren entscheidet. Für das Aufbauverfahren spräche wohl der anfangs geringere Rechenaufwand, während das Abbauverfahren vom vollen „Informationsgehalt“ der Daten ausginge, d. h. den Messwerten für alle fünf Wellenlängen. Wir vertiefen derartige Erörterungen nicht, sondern probieren beide Verfahren unvoreingenommen aus.

#### (2.2.1) Abbauverfahren

Für alle drei Datensätze werden die Wellenlängen 290 und 312 nm als entbehrlich erkannt. Bei deutlichem Rauschen kann auch auf die Wellenlänge 544 nm verzichtet werden.

Im Falle des stärksten Rauschens ist dann eher die Wellenlänge 526 nm entbehrlich.

Mit anderen Worten: In jedem Fall bleiben *zwei Wellenlängen* übrig, mit den man die neun Kalibrierklassen signifikant unterscheiden kann. Bei kaum verrauschten Daten sind dies die Wellenlängen 350 und 544 nm.

Das ist insofern eine Überraschung, als die von uns im Zweikomponenten-Teil bevorzugte Wellenlängen-Kombination (350 und 526 nm) sich ab stärkerem Rauschen als trennschärfste Variablen-Kombination erweist.

### (2.2.2) Aufbauverfahren

Das Aufbauverfahren führt immer zu Wellenlängen-Mengen, die die Wellenlängen 350, 526 und 544 nm enthalten. Bei gering verrauschten Daten treten noch die Wellenlängen 290 und 312 nm hinzu, bei mittlerem Rauschen nur die Wellenlänge 312 nm und im stark verrauschten Fall nur die Wellenlänge 290 nm.

Die entsprechenden numerischen Werte der kanonischen Variablen etc. können wir hier übergehen. Mit diesen könnten wir zwar weitere (neue, unbekannte) Proben den bekannten neun Kalibrierklassen zuordnen, was uns an dieser Stelle jedoch nicht interessiert. In der Projektion der Kalibrierversuche in die Ebene der ersten beiden kanonischen Variablen  $Y_{KV}$  (*Diskriminanzfunktionen*) ist aber schön zu sehen, wie die Klassenstruktur zunehmend verschwommener wird.

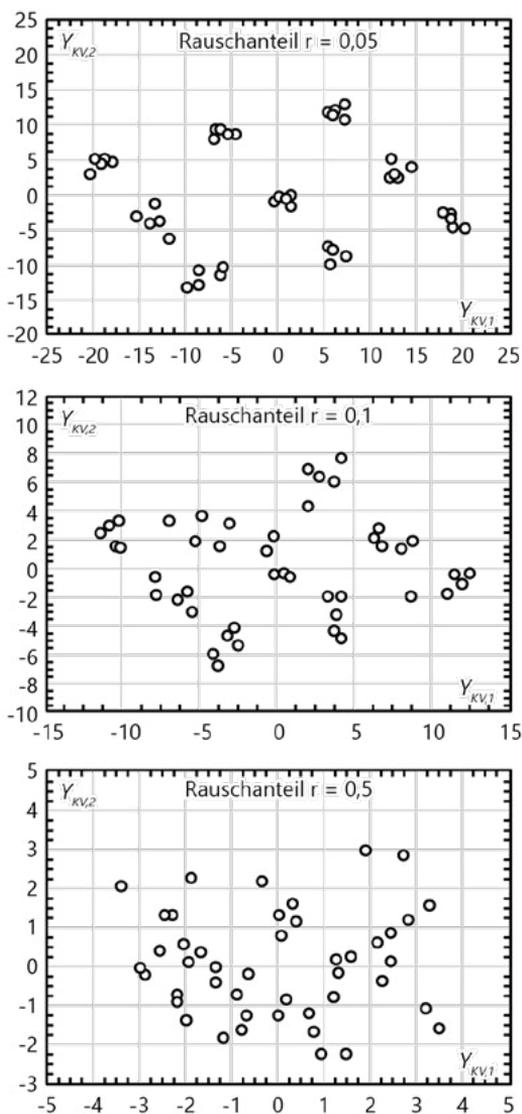


Abb. 22: Diskriminanzanalyse: 3 Projektionen der Kalibrierversuche bei zunehmendem Rauschanteil  $r$  der Spektren

## 6.4.7 DATENBEISPIEL ZU 6.3.3 - „SYSTEMISCHE“ SIGNAL/RAUSCH-MODELLIERUNG (LDA)

Vor Jahrzehnten hatte ZWANZIGER vorgeschlagen, im Sinne einer „Systemischen“ Optimierung das Signal-Rausch-Verhältnis von Analysemethoden mittels der Diskriminanzanalyse (LDA) zu verbessern. Ob dies heute noch eine lohnende Aufgabe ist, mag der interessierte Leser entscheiden. Es könnte immerhin sein, dass die Messtechnik unterdessen kaum noch zufällig verrauschte Signale misst. Auch sollten etwa einem „Dosierroboter“ kaum noch nennenswerte Zufallsfehler unterlaufen. Von systematischen oder handwerklichen Fehlern im Versuchsdesign sehen wir gänzlich ab.

Um die Erfolgsaussichten des Optimierungsvorschlags zu testen, verwenden wir in diesem Beispiel die LDA ohne die Absicht einer Variablenreduktion. Vielmehr sollen die Diskriminanzfunktionen die Beiträge aller erfassten Wellenlängen enthalten.

- ❶ Daten-ID: 21.01.2020; Mischungsproben, Design:  $3^2$ -VP, 5 Wellenlängen; Einfachmessungen; daraus simulierte Mehrfachmessungen; Konzentrationen in  $\text{mmol l}^{-1}$

Mit den Daten des Zweikomponenten-Falls, die wir bereits mehrfach verwendet haben, wollen wir nun systematische Modellierungsversuche unternehmen und deren Empfindlichkeitsmatrizen miteinander vergleichen.

Separate (sequentielle) Einkomponenten-Modelle würden eine extrem gestörte Kalibrationsaufgabe beschreiben, denn faktisch liegt neben dem jeweiligen Analyten (Cr oder Mn) gleichzeitig eine Störkomponente in Form eines zweiten Analyten vor (Mn oder Cr). Als Datensatz für unsere Software (STATISTICA) verwenden wir die Einfachmessungen aus dem  $3^2$ -Versuchsplan.

Der „Informations-Block“, an den sich rechts die Signalmatrix anschließt, sieht so aus:

Probe- name	Mn- Konzentration	Cr- Konzentration	mögl. Klassen		
			$X_1$	$X_2$	VP
M11	0,08	0,15	1	1	11
M12	0,08	0,30	1	2	12
M13	0,08	0,45	1	3	13
M21	0,20	0,15	2	1	21
M22	0,20	0,30	2	2	22
M23	0,20	0,45	2	3	23
M31	0,32	0,15	3	1	31
M32	0,32	0,30	3	2	32
M33	0,32	0,45	3	3	33
Fallunterscheidung:			3K	3K	9K

Über die Klassifikationsvariablen können mehrere Fälle eindeutig unterschieden werden: für Mn und für Cr je drei Konzentrationsklassen (Dreipunkt-Kalibrationsansätze) und ein VP-Fall mit 9 VP-Klassen, den wir mittels LDA modellieren werden.

Bei den Dreiklassen-Fällen wäre jeweils ein Einkomponenten-Kalibrationsmodell oder ein Zweikomponenten-Modell denkbar,  $Y(\lambda_p) = f(\text{Mn}, \text{Cr})$ .

Da beide Analyt-Konzentrationen zur Kalibration bekannt sind, werden wir diese Information auch nutzen und uns nur der Zweikomponenten-Kalibration im Fall mit neun Klassen zuwenden.

Zur Ermittlung der Modell-Koeffizienten bedienen wir uns wiederum der multiplen Regressionsanalyse (MLR). Obwohl eine *multivariate* Sachlage ( $P = 5$ ) vorliegt und wir diese unten auch entsprechend mittels Matrix-Algebra beschreiben werden, ermitteln wir die Modell-Koeffizienten für jede Wellenlänge separat (sequentiell).

Die multiple lineare Regression wird im Folgenden einerseits auf die *Originaldaten* angewendet, als auch auf *transformierte Daten*, die mittels Diskriminanzanalyse (LDA) gewonnen worden sind.

Man beachte den Unterschied: Während die Original-Variablen  $Y_{\text{ORIG}}$  die „Klassen-Information“ nicht enthalten, sollten diese in den Diskriminanz-Variablen  $Y_{\text{DV}}$  latent enthalten sein. Wir könnten deshalb in Analogie zur PCR auch von „LDR“ sprechen (linear discriminant regression).

## (1) Die Daten

Die Signalwerte der Einfachmessungen wurden zunächst mehrfach kopiert, so dass jede Versuchsklasse (Faktorkombination) fünffach besetzt<sup>83</sup> ist und sich  $N = 45$  ergibt.

Zusätzlich zum Block der fünf Original-Messgrößen (Variablen; Wellenlängen) haben wir dann vier weitere Blöcke (Datensätze) mit zunehmend verrauschten Messwerten erzeugt, indem zu den Ausgangswerten normalverteilte Zufallswerte aus einem bestimmten Intervall  $\{0 | r\}$  addiert wurden. (Die Rauschamplitude kürzen wir mit „ $r$ “ ab.)

Somit können wir nun zusätzlich mehrere „Schwierigkeitsgrade“ erzeugen und folgende Datensätze miteinander vergleichen:

Datensatz „ $r = 0,005$ “: Daten „kaum“ verrauscht,

Datensatz „ $r = 0,05$ “: Daten „mäßig“ verrauscht,

Datensatz „ $r = 0,1$ “: Daten „stark“ verrauscht,

Datensatz „ $r = 0,5$ “: Daten „extrem“ verrauscht.

Die numerischen Werte in den Verfahrensschritten für die beiden Varianten (mit Original- und mit transformierten Variablen) beziehen sich exemplarisch auf den Datensatz „ $r = 0,05$ “. In den Fehlertabellen stellen wir Ergebnisse für alle Datensätze gegenüber.

## (2) Verfahrensschritte mit Originalvariablen

### (2.1) Festlegen der Modellgleichung

Die Modellgleichung sei:

$$Y(\lambda_p) = b_2 X_2 + b_1 X_1 + b_0,$$

wobei die ungewohnte Reihenfolge lediglich einen programmtechnischen Grund hat: die verwendete EXCEL-Funktion `rgp(...)` liefert die Koeffizienten in dieser Reihenfolge. Mit den leerwertkorrigierten Signalwerten  $Y'$  lautet die Modellgleichung:

$$Y'(\lambda_p) = Y(\lambda_p) - b_0 = b_1 X_1 + b_2 X_2,$$

mit  $p = 1, \dots, P$ ;  $P = 5$ ;  $N = 45$ ;  $C = 2$ ;  $X_1 = \text{Mn}$ ,  $X_2 = \text{Cr}$ .

In Matrix-Schreibweise ist unser Modell dadurch:

$\mathbf{Y}' = \mathbf{X} \mathbf{B}$ , mit den konkreten Matrix-Dimensionen (45,5), (45,2) und (2,5).

---

<sup>83</sup> Untersuchungen zum Stichprobenumfang bei der Konstruktion linearer Klassifikatoren haben ein Verhältnis von  $N/P > 3$  gefordert, mitunter sogar für jede Klasse  $k$  verlangt:  $N_k/P > 3$  (MASSART, pers. Mitteilung, 1984). Eine Variablenreduktion ist also manchmal ratsam, wenn man die Versuchsanzahl nicht erhöhen kann.

(2.2) Überprüfen der Kondition der VP-Matrix

Dazu bilden wir (mit allen Konzentrationswerten):

$$\mathbf{X}^T \mathbf{X} = \begin{pmatrix} 2,232 & 2,700 \\ 2,700 & 4,725 \end{pmatrix}$$

Die Determinante dieses Matrixprodukts (Wert = 3,26) deutet an, dass der MLR-Algorithmus bei der Inversion von  $\mathbf{X}^T \mathbf{X}$  keine Probleme haben wird. (Interessant, dass sich der Determinantenwert der „1. Wiederholung“ des Versuchsplans mit der Zahl der Wiederholungen des Versuchsplans vervielfacht und damit weiter von Null entfernt.)

(2.3) Berechnen der Koeffizienten von  $\mathbf{B}_{\text{ORIG}}$  aus  $\mathbf{Y}_{\text{ORIG}}$

Für den jeweiligen Datensatz ermitteln wir nun die Elemente der  $(P,C+1)$ -Matrix  $(\mathbf{B}^*)^T_{\text{ORIG}}$  mit der Dimension (5,3).

Aus ihr bleibt nach Entfernung der Leerwertspalte die (5,2)-Matrix  $\mathbf{B}^T_{\text{ORIG}}$  übrig.

$\lambda$	$\mathbf{b}^T_{\text{Mn}}$	$\mathbf{b}^T_{\text{Cr}}$	$\mathbf{b}^T_0$	X	$\mathbf{B}_{\text{ORIG}} =$
290	<b>1,075</b>	<b>2,549</b>	-0,021	Mn	<b>1,075</b> <b>1,787</b> <b>1,245</b> <b>2,314</b> <b>2,212</b>
312	<b>1,787</b>	<b>1,520</b>	-0,052	Cr	<b>2,549</b> <b>1,520</b> <b>3,318</b> <b>-0,007</b> <b>-0,056</b>
350	<b>1,245</b>	<b>3,318</b>	-0,038		
526	<b>2,314</b>	<b>-0,007</b>	0,031		
544	<b>2,212</b>	<b>-0,056</b>	0,015		

(2.4) Leerwert-Korrektur ausführen

Die Leerwertkorrektur führt  $\mathbf{Y}_{\text{ORIG}}$  in  $\mathbf{Y}'_{\text{ORIG}}$  über.

(2.5) Anwenden des Modells

Die erste Anwendung erfolgt gemäß

$$\mathbf{X}_{\text{PRED}} = \mathbf{Y}_{\text{ANW}} \mathbf{A}_{\text{ORIG}}$$

auf die „Lerndaten“

$$\mathbf{Y}_{\text{ANW}} = \mathbf{Y}'_{\text{ORIG}}$$

Die Anwendungsmatrix („Analysenmatrix“)

$$\mathbf{A}_{\text{ORIG}} = \mathbf{B}^T_{\text{ORIG}} (\mathbf{B}_{\text{ORIG}} \mathbf{B}^T_{\text{ORIG}})^{-1}$$

hat folgende Werte:

	Mn	Cr
$(\lambda_1 = 290 \text{ nm})$	-0,012	0,134
$(\lambda_2 = 312 \text{ nm})$	0,091	0,033
$(\lambda_3 = 350 \text{ nm})$	-0,029	0,181
$(\lambda_4 = 526 \text{ nm})$	0,199	-0,095
$(\lambda_5 = 544 \text{ nm})$	0,192	-0,095

## (2.6) Fehlerberechnung

Man könnte zunächst (nur zur Orientierung) die *absolute prozentuale Abweichung* der berechneten  $X_{\text{PRED}}$  von den bekannten Konzentrationswerten errechnen. Diese Werte, gemittelt über die (willkürlich ausgewählten) ersten 9 Versuchspunkte, vermitteln einen ersten Eindruck von der Brauchbarkeit des Modells.

Die für uns aussagekräftigsten Kenngrößen sind jedoch wieder *SEC* bzw. *SEP*, die wir bereits verwendet haben. Wie auch der Gesamtfehler *TE* lassen sie sich aus dem Wert von *PRESS* ableiten (vgl. Gleichung 58). Aus numerischen Gründen tabellieren wir den Gesamtfehler *TE*, den wir anhand aller 45 Versuchspunkte errechnet haben.

Datensatz mit „Rauschpegel“ $r =$	abs. Abweichungen (in %)		$TE_{\text{ORIG}}$	
	Mangan	Chrom	Mangan	Chrom
0,005	0,5	0,5	0,01	0,01
0,05	4,7	3,1	0,09	0,08
0,1	12,6	5,5	0,18	0,18
0,5	8,1	31,1	0,81	0,53

Vorläufiges Fazit:

- Die mit dem Rauschanteil zunehmende Fehlerquote wird besonders bei den mittleren prozentualen Abweichungen deutlich.
- Anhand von *TE* erkennt man jedoch eine im Mittel erstaunliche Toleranz des Rauschanteils durch die MLR-Methode.
- Wir interpretieren das so: Der gute (statistische) Versuchsplan erweist sich erneut als vorteilhaft.

## (3) Verfahrensschritte mit Diskriminanzvariablen

### (3.1) Ausführen der Diskriminanzanalyse

Mit  $K = 9$  Konzentrationsklassen des  $3^2$ -Versuchsplans und  $P = 5$  Wellenlängen gibt es nach  $T = \min(K - 1, P)$  genau  $T = 5$  von Null verschiedene Eigenwerte mit ihren zugehörigen Eigenvektoren.

Mit den Eigenvektoren lassen sich prinzipiell die Original-Variablen  $Y_{\text{ORIG}}$  in Diskriminanz-Variablen  $Y_{\text{DV}}$  transformieren. Das leider nicht auf „Kalibrierungsaufgaben“ spezialisierte Programmpaket STATISTICA gibt *einerseits* direkt „unstandardisierte kanonische Werte“ für jeden Versuchsfall aus. Sie bilden die Werte-Matrix  $\mathbf{Y}_{\text{KW}}$ . Diese Werte der Versuchsfälle lassen sich in einem ebenen Streudiagramm der beiden kanonischen Variablen darstellen und illustrieren anschaulich die mit zunehmendem Rauschanteil fortschreitende Vermischung der Versuchsklassen (Analytkombinationen). *Andererseits* können wir die Matrix  $\mathbf{U}$  der „unstandardisierten Koeffizienten für kanonische Variable“ selber anwenden, um eine beliebige Messwerte-Matrix  $\mathbf{Y}$  in eine Matrix  $\mathbf{Y}_{\text{TC}}$  zu transformieren, die mit  $\mathbf{Y}_{\text{KW}}$  identisch ist.

Eine Variante zur Validierung könnte dann beispielsweise wie folgt ablaufen: Die mittels einer „worst case“-Matrix  $\mathbf{Y}_{\text{KAL}}$  (z. B.  $\mathbf{Y}_{\text{ORIG}}$  des Datensatzes „ $r = 0,5$ “) ermittelte „Anwen-

dungsmatrix"  $\mathbf{A}$  wird im Validierschritt auf eine „unproblematische“ Matrix  $\mathbf{Y}_{ANW}$  (z. B.  $\mathbf{Y}_{ORIG}$  des Datensatzes „r = 0,05“) angewandt oder umgekehrt. (Vergleiche dazu die dritte und vierte  $TE$ -Spalte in der Tabelle unten.)

### (3.2) Transformieren der Original-Daten

Die von STATISTICA ausgegebene  $(P+1, T)$ -Matrix  $\mathbf{U}$  enthält neben den „unstandardisierten Koeffizienten“ noch einen Zeilenvektor **const** mit Konstanten für jede latente Variable  $t, t = 1, \dots, T$ . Die Spaltenvektoren  $\mathbf{u}$  von  $\mathbf{U}$  sind die „Gewichte“, mit denen die  $P$  Originalvariablen in jede Diskriminanzvariable  $\mathbf{Y}_t$  eingehen. Für jeden unserer  $N$  Versuche ergeben sich somit die Werte seiner kanonischen Variablen gemäß

$$\mathbf{y}_{TC} = \mathbf{const} + \mathbf{y}_{ORIG} \mathbf{u} \text{ mit den Dimensionen:}$$

$$(1, T) \quad (1, T) \quad (1, P) \quad (P, T).$$

### (3.3) MLR mit transformierten Messwerten

Von hier an verfahren wir genau wie oben ab Schritt (2.3).

Auf einige Unterschiede ist aber hinzuweisen.

Bei der MLR mit Originalvariablen erhielten wir zunächst eine  $(P, C+1)$ -Matrix  $(\mathbf{B}^*)^T_{ORIG}$  mit  $P$  Zeilen für jede Wellenlänge, nach Anwenden der LDA aber eine  $(T, C+1)$ -Matrix  $(\mathbf{B}^*)^T_{ORIG}$  mit  $T$  Zeilen. Jede dieser Zeilen entspricht einer Diskriminanzvariablen  $Y_{DV}$ , die ihrerseits eine Linearkombination der  $P$  Originalvariablen (5 Wellenlängen) ist.

Deren Koeffizienten wurden nur unter Beachtung der Klassenstruktur berechnet.

Für die MLR ist der Regressor  $\{X_1, X_2\}$  in beiden Verfahrensfällen gleich, der Regressand unterscheidet sich jedoch. Die Elemente der Empfindlichkeiten-Matrizen  $\mathbf{B}_{ORIG}$  und  $\mathbf{B}_{TC}$  werden aber in der gleichen Weise berechnet, wie oben geschildert.

Es ergeben sich die konkreten Werte für den Datensatz „r = 0,05“:

$\mathbf{B}_{TC} =$	$Y_{TC1}$	$Y_{TC2}$	$Y_{TC3}$	$Y_{TC4}$	$Y_{TC5}$
Mn	49	-68	0,08	0,04	0,05
Cr	87	25	- 0,04	0,02	0,02

Daran erkennt man übrigens deutlich, dass eigentlich nur die ersten beiden Diskriminanzvariablen von Bedeutung sind, woraus auch die „verlustlose“ ebene Darstellung resultiert.

### (3.4) Leerwert-Korrektur ausführen

Die Leerwertkorrektur führt  $\mathbf{Y}_{TC}$  in  $\mathbf{Y}'_{TC}$  über.

### (3.5) Anwenden des Modells

Die Anwendung erfolgt formal wie bekannt, nur mit alternativen Empfindlichkeiten-Matrizen.

Die Anwendungsmatrix  $\mathbf{A}_{TC} = \mathbf{B}_{TC}^T (\mathbf{B}_{TC} \mathbf{B}_{TC}^T)^{-1}$  hat für den Datensatz „r = 0,05“ folgende Werte:

$\lambda$	Mn	Cr
290	0,003	0,010
312	-0,012	0,007
350	0,000	0,000
526	0,000	0,000
544	0,000	0,000

### (3.6) Fehlerberechnung

Folgende Tabelle zeigt einige Ergebnisse der Anwendung der  $\mathbf{A}$ -Matrizen.

Die Kenngröße  $TE$  (45 Versuchspunkte) ist proportional der Fehlergröße  $SEC$  (Anwendung auf  $\mathbf{Y}_{TC}$ ) bzw. proportional der Fehlergröße  $SEP$  (Anwendung auf  $\mathbf{Y}_{ORIG}$ ).

In der Spalte  $TE$  ( $VAL_1$ ) sind die Fehler bei Anwendung von  $\mathbf{A}$ , ermittelt mit Datensatz  $r = 0,5$ , auf den Datensatz  $r = 0,005$  angegeben und in Spalte  $TE$  ( $VAL_2$ ) die Fehler bei Anwendung von  $\mathbf{A}$ , ermittelt mit Datensatz  $r = 0,005$ , auf den Datensatz  $r = 0,5$ .

Datensatz mit „Rauschpegel“ r =	$TE$ ( $SEC$ )		$TE$ ( $SEP$ )		$TE$ ( $VAL_1$ )		$TE$ ( $VAL_2$ )	
	Mn	Cr	Mn	Cr	Mn	Cr	Mn	Cr
0,005	0,01	0,01	3,6	3,5	3,1	4,0		
0,05	0,09	0,08	3,7	3,2				
0,1	0,17	0,17	3,8	3,5				
0,5	0,79	0,48	5,2	5,3			5,7	4,8

Fazit:

- Bei der Klassifikation, die nicht Gegenstand der Erörterung sind, traten auch im extrem verrauschten Datensatz keine Klassifikationsfehler auf.
- Wiederum verblüffen die erstaunlich niedrigen Gesamtfehler  $TE$ . Geteilt durch die Wurzel aus den jeweiligen Freiheitsgraden, ergibt sich natürlich das gleiche Muster für  $SEC$  und  $SEP$ .
- Spalten 4 und 5 deuten auf eine gute „Übertragbarkeit“ von Empfindlichkeitenmatrizen zwischen Kalibriermodellen hin. Es ist plausibel, dass eine mit einem „harmlosen Fall angerlegte“ Empfindlichkeitenmatrix im „Härtefall“ (Wertepaar rechts unten) schlechtere Ergebnisse bringt als umgekehrt.
- Insgesamt hat sich gezeigt, dass die LDA ebenso niedrige Fehler wie die traditionelle Verfahrensweise erzielt. Das ist nicht verwunderlich, denn sie arbeitet mit dem gleichen Informationsgehalt, den die Original-Messvariablen haben, die sie „lediglich gut kombiniert“. Es muss weiteren Untersuchungen vorbehalten bleiben, ob sie bei schlechterer Versuchsplanung (zufällige Kalibration) Vorteile bringt.

## 7 Literatur

---

### 7.1 Einkomponentenanalytik

R. CAULCUTT, R. BODDY: „Statistics for Analytical Chemists“, Chapman and Hall, London, 1983

K. DOERFFEL: „Statistik in der analytischen Chemie“, 5. erw. u. überarb. Aufl., Dt. Verlag für Grundstoffindustrie, Leipzig, 1990

ISO Standards Handbook 3, International Organization for Standardization, 3rd ed., Genf, 1989

H. KAISER: „Guiding concepts relating to trace analysis“, Pure & Appl. Chem. 34(1973)35-61

H. MAGER: „Moderne Regressionsanalyse“, Salle und Sauerländer, Frankfurt, Berlin, München, Salzburg, Aarau, 1982

J. N. MILLER: „Calibration Methods in Spectroscopy. I.“ Spectrosc. Int. 3(1991) no. 4, 41-43

J. C. MILLER, J. N. MILLER: „Statistics for Analytical Chemistry“, 2nd ed., Ellis Horwood Series in Analytical Chemistry, Ellis Horwood, Chichester, 1988; 3rd ed., Ellis Horwood, Chichester, 1994

P. H. MÜLLER, P. NEUMANN, R. STORM: „Tafeln der mathematischen Statistik“, 3. Aufl., Fachbuchverlag Leipzig, Leipzig, 1979

P. J. ROUSSEEUW, A. M. LEROY: „Robust Regression and Outlier Detection“, Wiley, New York, 1987

R. STORM: „Wahrscheinlichkeitsrechnung, Mathematische Statistik, Statistische Qualitätskontrolle“, 10., völlig Neubearb. Aufl., Fachbuchverlag Leipzig, Köln, 1995

### 7.2 Mehrkomponentenanalytik

K. R. BEEBE, B. R. KOWALSKI: „An Introduction to Multivariate Calibration and Analysis“, Anal. Chem. 59(1987)1007A-1017A

C. W. BROWN, R. J. OBREMSKI: „Multicomponent Quantitative Analysis“, Appl. Spectrosc. Rev. 20(1984)373-418

H. KAISER: „Zur Definition von Selektivität, Spezifität und Empfindlichkeit von Analysenverfahren“, Z. Anal. Chem. 260(1972)352-260

H. MARTENS, T. NAES: „Multivariate calibration. I. Concepts and distinctions“, Trends Anal. Chem. 3(1984)204-210

H. MARTENS, T. NAES: „Multivariate Calibration“, Wiley, Chichester, 1989

L. D. MASSART, B. G. M. VANDEGINSTE, S. N. DEMING, Y. MICHOTTE, L. KAUFMAN: „Chemometrics: a textbook“, Elsevier, Amsterdam, 1988

T. NAES: „Comparison of Approaches to Multivariate Linear Calibration“, Biom. J. 27(1985)265-275

T. NAES, H. MARTENS: „Multivariate calibration. II. Chemometric methods“, Trends Anal. Chem. 3(1984)266-271

M. A. SHARAF, D. L. ILLMAN, B. R. KOWALSKI: „Chemometrics“, Wiley, New York, 1986, chapter 4: „Calibration and Chemical Analysis“

## Anhang: Orientierungshilfe zur Methodenauswahl für Multivariate Datenauswertung und Modellierung

Diese Orientierungshilfe umfasst alle Methoden, die zu den „kanonischen Methoden“ der interdisziplinären „Chemometrik“ zählen. Will man diese Methoden „mathematischen Laien“ nahebringen, muss eingangs betont werden: Datenanalyse (Chemometrik) ist einfacher als man denkt – wenn man nur genau weiß, was man eigentlich wissen will bzw. welche Information man braucht.

Man verfährt deshalb zweckmäßig wie folgt:

(1) **Ordnen der Daten,**

d. h. der Werte, von denen die Datenanalyse ausgehen soll

Dabei ist zu überlegen, für welche „Objekte“ man sich interessiert und für welche „Merkmale“ (Eigenschaften, genauer: Eigenschaftsgrößen oder Messgrößen) dieser Objekte Zahlenwerte vorhanden sind.

*Beispiele für Objekte:* Versuche, Proben, Spektren, Zeitpunkte, Geräte, Leute etc.

*Beispiele für Merkmale:* Konzentrations- und Intensitätsgrößen, Qualitätskriterien, Stoffgrößen, Versuchsparameter etc.

In Tabellenform werden zeilenweise die Objekte und deren Merkmale angeordnet.

In der ersten Spalte stehen die Bezeichnungen (Namen) der Objekte und in den weiteren Spalten die (Zahlen-)Werte der Merkmale.

*Beispiel: Wasserproben*

Objekt	Merkmale		
	Geschmack	Härte (°dH)	Cd ( $\mu\text{g l}^{-1}$ )
Probe 1	gut	12	3
Probe 2	schlecht	6	2
Probe 3	schrecklich	8	20

„Block O“

„Block X“

Der „Block X“ enthält also die *Rohdaten*, er bildet die eigentliche *Datenmatrix X*.

Wenn man nichts weiter aufschreiben will oder kann, wendet man sich nun Punkt (3) zu, sonst Punkt (2).

(2) **Erweitern der Daten-Tabelle,**

sofern man die Objekte noch genauer beschreiben kann.

Damit bezieht man zusätzliches Wissen (nichtanalytisches „Vorwissen“) in die Datenanalyse ein.

Dazu werden einfach entsprechend viele Spalten in die Tabelle eingefügt.

Meist kann man mit solcher Zusatzinformation die Daten bereits sinnvoll klassifizieren.

Wir nennen die Spalten deshalb auch Klassifikations- oder Indexvariablen („Block K“).

Beispiel: Wasserproben

Objekt (Name)	Klassifikation (Zuordnung)		Merkmale		
	Herkunft	Art	Geschmack	Härte	Cd
Campus	MERSEBURG	LEITUNG	gut	122	3
Fam. Müller	MERSEBURG	LEITUNG	gut	231	4
Fam. Meyer	MERSEBURG	BRUNNEN	schlecht	62	2
Fam. Schulz	LEIPZIG	BRUNNEN	schrecklich	84	23
T. A.	LEIPZIG	REGEN	gut	43	2
Südfriedhof	LEIPZIG	REGEN	gut	50	1,9

„Block O“    „Block K“                    „Block X“

(3)    **Formulieren konkreter Fragen**  
oder Hypothesen zu den Daten

Man sollte immer zuerst überlegen, ob nicht bereits einfache graphische Darstellungen wie Häufigkeits- und Streudiagramme oder Zeitreihendarstellungen („Kontrollkarten“)<sup>84</sup> bei der Problemlösung helfen.

Viele Auffälligkeiten, Trends etc. werden bei solchen Darstellungen schon deutlich und helfen, tiefergehende Fragen bzw. Hypothesen zu formulieren.

Zur richtigen Auswahl der passenden Basismethoden der Datenanalyse orientiert man sich am besten wie folgt und anhand der unten stehenden tabellarischen Zusammenfassung.

(3.1)

Wenn das Problem oder die Fragen vorwiegend mit „Block O“ bzw. „Block O“ und „Block X“ zu tun haben, helfen wahrscheinlich schon die Methoden Clusteranalyse (CA) und Hauptkomponenten- bzw. Faktoranalyse (PCA bzw. FA) weiter.

*Beispiele für konkrete Fragen:*

- Welche Gruppierungen ähnlicher Objekte gibt es? (→ CA, auch PCA)
- Welche Merkmale verhalten sich ähnlich? (→ PCA, FA, auch CA)

(3.2)

Beziehen sich die Fragen auf „Block K“ bzw. „Block K“ und „Block X“, ist wahrscheinlich die Varianz- und Diskriminanzanalyse (VA, DA) die Methode der Wahl.

*Beispiele für konkrete Fragen:*

- Lassen sich Gruppen von Objekten statistisch unterscheiden? (→ DA)
- Lässt sich diese Unterscheidung bei gleicher statistischer Sicherheit mit weniger Aufwand (weniger Merkmalen) bewerkstelligen? (→ VA und DA)
- Lässt sich ein „unbekanntes“ Objekt einer Gruppe mit großer Wahrscheinlichkeit zuordnen? (→ DA)

---

84 Unter diesem Begriff wird man im Programmpaket STATISTICA brauchbare Methoden finden.

(3.3)

Interessiert man sich für (qualitative) Zusammenhänge der Merkmale (Größen) in „Block X“, so können wieder Cluster- sowie Hauptkomponenten- oder Faktoranalyse nützlich sein. Maßzahlen für diese Zusammenhänge liefert die Korrelationsanalyse.

Wechselseitige Abhängigkeiten von Größen aus „Block X“ werden (quantitativ) durch Methoden der einfachen oder multivariaten (multiplen) Regressionsanalyse (RA oder MLR) beschrieben.

*Beispiele für konkrete Fragen:*

- Besteht ein Zusammenhang zwischen den Merkmalen 2 und 7? (→ Korrelationskoeffizient)
- Welche funktionelle Abhängigkeit des Merkmals 3 vom Merkmal 8 besteht? (→ Regressionsmodelle)
- Von welchen Einflussgrößen hängt eine Zielgröße (signifikant) ab? (→ multivariate Regressionsanalyse)

Wenn sich die Fragestellung keiner der o. a. Basismethoden CA, PCA (FA), RA (MLR) und VA (DA) zuordnen lässt oder wenn man unsicher ist, fragt man am besten einen ausgeprägt anwendungsorientierten Mathematiker (Datenanalytiker) nach einer geeigneten Methode.

(4) **Auswertung der Daten**

mittels einer Software, die die o. a. Methoden enthält

Für eine optimale Nutzung der Software, die in der Regel zahlreiche Feineinstellungen erfordert, und zur sachlichen Interpretation und glaubhaften Verwendung der Ergebnisse hat man aber am besten einen Datenanalytiker an seiner Seite.

(5) **Software-Empfehlung**

(für Anspruchsvolle bzw. Fortgeschrittene)

Datenerfassung ggf. mittels EXCEL<sup>®</sup> (Microsoft Corporation) vornehmen und EXCEL-Dateien in einem geeigneten Format exportieren.

So kann ein Format gewählt werden, um zum Beispiel nachfolgend Dateneingabe, Graphikerstellung und Auswertungen mit STATISTICA<sup>®</sup> (StatSoft Inc.), vorzunehmen. Dies ist bei großen Block-X-Tabellen empfehlenswert. Spezielle, auf Spektrenauswertung zielende Software wie UNSCRAMBLER<sup>®</sup> (CAMO Analytics), kann für große Block-Y-Tabellen (siehe Punkt (7)) hilfreich sein. Heutzutage liefern die marktbestimmenden Firmen zu ihrer Analysetechnik aber in der Regel angepasste Software mit aus.

(6) **Methoden-Orientierung**

bei der Auswahl chemometrischer Basis-Methoden mittels folgender Zusammenfassung

	Interesse an <b>Zusammenhängen</b> , also an qualitativen Beziehungen:	Interesse an <b>Abhängigkeiten</b> , also an quantitativen Beziehungen:
	<b>Ohne</b> Verwendung von <b>Zusatzinformation</b> zu den Daten:	<b>Mit Zusatzinformation</b> zu den Daten, die z. B. eine bestimmte <b>Gruppierung</b> ermöglicht:
Bei Fragen, die vor allem die <b>Objekte</b> betreffen:	CA, PCA, FA	DA
Bei Fragen, die vor allem die <b>Merkmale</b> betreffen:	PCA, FA, CA CCA	VA, RA, MLR

Die Abkürzungen bezeichnen folgende „Methodenfamilien“:

CA Clusteranalyse  
 CCA Korrelationsanalyse oder Kanonische Korrelation  
 DA Diskriminanzanalyse (lineare)  
 FA Faktoranalyse  
 MLR multivariate (multiple) lineare Regressionsanalyse  
 PCA Hauptkomponentenanalyse  
 RA Regressionsanalyse  
 VA Varianzanalyse

(7) **Multivariate Kalibration (Multivariate Modellierung)**

Die Datentabellen bestehen vor allem oder ausschließlich aus einem umfangreichen „Y-Block“.

*Beispiel:*

Objekt (Name)	Klassifikation		Berechnete oder gemessene Merkmale		
	Wasser (von)	Tiefe (in cm)	Härte (in °d)	Cd (in µg l <sup>-1</sup> )	Spektrum mit Intensitäten bei $\lambda_1, \dots, \lambda_p, \dots, \lambda_p$
Probe 1	Ort 1	20	122	3	<P Messwerte>
Probe 2	Ort 2	40	231	4	<P Messwerte>
...	...	...	...	...	
Probe n	Ort 3	30	84	23	<P Messwerte>
...	...	...	...	...	
Probe N	Ort 728	35	50	1,9	<P Messwerte>

„Block O“    „Block K“                    „Block X“                                    „Block Y“

Nun müssen wir uns nur noch klarmachen, dass die unterschiedlichen datenanalytischen Methoden sowohl auf Werte des „Blocks X“ als auch auf Werte des „Blocks Y“ getrennt angewendet werden können oder auf beide Blöcke gleichzeitig, und zwar jeweils mit oder ohne Einbeziehung der Information aus „Block K“.

Die Probenbezeichnung in „Block O“ sollte einer gewissen Systematik unterliegen, um später Objekte eindeutig identifizieren zu können.

## Verzeichnis der Abbildungen

---

Abb. 1: Linearer Kalibriergraph mit Konfidenzband	44
Abb. 2: Q-Q-Plot (unten für $X$ , oben für $Y$ )	46
Abb. 3: Typische heteroskedastische Datenlage	46
Abb. 4: Auswirkung eines Ausreißers im $Y$ -Bereich	47
Abb. 5: Auswirkung eines Ausreißers in $X$ -Bereich	47
Abb. 6: Normalverteilungs-Plot der Residuen	48
Abb. 7: ANSCOMBE-Quartett	48
Abb. 8: Signalwerte mit Fehlerbalken	53
Abb. 9: Kalibriergraph mit Konfidenzband	54
Abb. 10: Illustration von Grenzwerten und Entscheidungsrisiken anhand der Häufigkeitsverteilungen ( $H$ ) definierter Signalwerte	63
Abb. 11: Einfache Analytzugabe	75
Abb. 12: UV-Spektren von Permanganat und Dichromat	100
Abb. 13: Univariate Kalibrierfläche $Y = f(\text{Mn}, \text{Cr})$	102
Abb. 14: Abbildung der Versuchspunkte (suboptimales Wellenlängenpaar)	127
Abb. 15: Abbildung der Versuchspunkte (optimales Wellenlängenpaar)	127
Abb. 16: Dendrogramm der 9 Versuchspunkte	129
Abb. 17: Dendrogramm der Wellenlängen	129
Abb. 18: Screeplot zur Hauptkomponentenanalyse	130
Abb. 19: Projektion der Kalibrierversuche mittels Hauptfaktoren	131
Abb. 20: Projektion der Wellenlängen mittels Hauptfaktoren	132
Abb. 21: Faktorladungen (unrotiert) der Faktoranalyse	135
Abb. 22: Diskriminanzanalyse: 3 Projektionen der Kalibrierversuche bei zunehmendem Rauschanteil $r$ der Spektren	137

## Sachwortverzeichnis

100%-Methode	71
AAS	29
Abhängigkeit, funktionale	19, 20
Abhängigkeiten	149
von Variablen	148
Absolutmethoden	17
Abstandsmaße	128
Abweichungen, quadratischen	20
Achsenabschnitt	20
Adäquatheitstest	33
AES	74
Analysenergebnisse [test results]	59
Analysenergebnisse, Vertrauensbereich	37
Analysenprobe	69
Analysenprobe [analytical portion]	65
Analysenprobe [laboratory sample, test sample]	17
Analyt	69
Analyt-Entfernung	81
Analyt-Interferenzen	116
Analyt-Subtraktion	81
Analyt-Zugabe	73
doppelte	79
einmalige	77
analyte addition	79
analyte subtraction	81
Analytgröße	69
ANOVA	34
Anpassungsdefekt, Test	33
Anpassungsgüte [goodness of fit]	32, 35, 52, 54
Test	32
Anstieg [slope]	19, 20, 24
Test	24
Vertrauensbereich	25
Anwendungsbeispiel	
Direktpotentiometrie	78
Polarographie	76
Potentiometrie	80
Aufstocken	17, 73
Ausreißer	115
Ausreißerererkennung	22, 115
visuelle	126

Ausreißertest	22, 49
Auswahlkriterien nach KAISER	104
Autoskalierung	127
BARTLETT-Test	55
Basisgrößen	16
Begriffe [notion]	61
Beispielrechnung	
Kalibrierung mit Einfachmessungen	44, 45, 49, 50, 52
Varianzanalyse	56
Vergleich von Regressionsgeraden	59
Bestimmtheitsmaß [coefficient of determination]	32, 35, 41, 43, 52
Test	32
Bestimmungsgrenze [limit of quantitation, determination]	66
blank value	20
Blindprobe	69
Blindwert	69
Bootstrap- Methode	40
Chemometrik	146
CHERNOFF-Gesichter	115
Chromatographie	83
Clusteranalyse	115, 127, 148, 149
COCHRAN-Test	21, 55
coefficient of determination	32
Darstellungen, graphische	147
Datenanalyse	146
Datenbeispiel	
2 <sup>2</sup> -Versuchsplan	103
3 <sup>2</sup> -Versuchsplan	103
Anwenden des Kalibriermodells	105
Direkte Kalibration	123
Inverse Kalibration	124
KAISERsche Auswahlkriterien	104
Kalibrierung mit Einfachmessungen (7 Messstellen)	44
Konditionszahl	106
Linear abhängige Kalibrierdaten	124
Maximalfehler	106
Modellierung, Modellbewertung	101, 103
Optimale Kalibrierproben	132
Regressionsfehler	103
Statistische Messstellenauswahl	134, 137
Visuelle Ausreißerkennung	126
Visuelle Spektrenbewertung	100
Datenmatrizen, dreidimensionale	115

Datenvorbehandlung	113
Datenvorbereitung, Hauptkomponentenzerlegung	113
Dendrogramm	115, 129
Determinantenkriterium	92
Diagonalmatrix	133
Direktpotentiometrie	73, 78, 81
Diskriminanzanalyse	114, 136, 137, 148, 149
lineare	119
Diskriminanzvariable	141
Distanzmatrix	128
DIXON-Test	21
Eichen	69
Eigenvektor	118
Eigenwert	118, 120
Eigenwertproblem	118, 120
Einheiten von Basisgrößen	16
Einpunkt-Kalibrierung	72
Einseitige [one-tailed]	64
EJCR-Methode	43
ELISA	86
Empfindlichkeit [sensitivity]	19, 20, 92
partielle	91
enzyme linked immuno sorbent assay	86
Erfassungsgrenze	66
EUKILDscher Abstand	128
EUKLID-Norm	107
Experimentalfehler, Berücksichtigung	50
externe Kalibrierung	70
externer Standard	70
Extinktion	17
Extremalproblem	120
Faktoranalyse	147, 149
Fehler	
systematischer	62
zufälliger	63
Fehler 1. Art [error of the first kind]	66
Fehlerquellen, systematische	62
Flächen-Korrekturfaktor	82
FROBENIUS-Norm	107
Fuzzy Regression	29
Gaschromatographie	71
Gehalt	16
generalized standard addition method	73

Gewicht [consequences, importance]	62
Gleichungssystem	
bestimmtes	91
überbestimmtes	91
GOF	32
goodness of fit	32
Größen	14
Analytgröße	14
Signalgröße	14
Größengleichung	19
GSAM	73
Gütemaße, Vergleich	42
Hauptachsentransformation	114
Hauptkomponentenanalyse	114, 115, 119, 130, 149
Hauptkomponentenregression	113
Hauptkomponentenzerlegung	97, 113, 114, 117
heteroskedastischer Fall	31
Homoskedastizität	55
HPLC	84
ICAP	74
Informationsmatrix	117, 133
Inkrementtechnik	73, 80, 81
Inner-Klassen-Streumatrix	136
intercept	20
Interferenzen	116
internal reference	82
internal standard	82
interne Kalibrierung	75, 82
interne Referenz	82
Interner Standard	82
IRLS	43
IRLS-Technik	29
jackknifing	40
K-Matrix-Methode	97
Kalibration	17
direkte	123
inverse	51, 124
nichtlineare	52
Kalibrierdaten, linear abhängige	124
Kalibrieren	19
Kalibrierfehler	41
Kalibrierfunktion	18

Kalibriermatrix	91
Determinante	92
Kalibriermodell	19
alternatives	24
nichtlineares	31
polynomiales	30
quasilineares	30
Kalibriermodelle	
alternative	50
Beurteilung	32
Kalibrierprobe [analytical calibration portions]	17, 69
Auswahl	113
optimal herstellen	116, 132
Kalibriersystem, Bewertung	92
Kalibrierung	17, 19
direkte	91
externe	70
fuzzy	43
indirekte	92
interne	70, 76, 82
inverse	72, 111
iterativ gewichtete	43
kontrollierte	92, 111
robuste	43
sequentielle	70
simultane	70
zufällige	92, 111
Kalibrierzugabe	73
Kanonische Korrelation	149
Kausalbeziehung	17
Kausalzusammenhang	20
Kenngrößen	
konzentrationsbezogene	103, 105
signalbezogene	45
Klassifikationsvariablen	146
kleinste Quadrate	20
known addition	73
known subtraction	80
Koeffizienten	
empirische Testwerte	24, 50
Statistik	50
Vertrauensbereiche	25, 50
Konditionszahl	98, 106, 115

Konfusion	67
Kontrollkarten	147
Konzentrationswechselwirkungen	116
Korrekturfaktor	84
Korrelationsanalyse	148, 149
Korrelationskoeffizient	32
Korrelationsmatrix	118
Kreuzvalidierung	40
Kreuzvalidierungstechniken	114
kritischer Wert	64
lack of fit	33
latente Variablen	109, 114
leaving one out	40
Leerprobe [blank analytical portion]	69
Leerprobe [blank]	17
Leerproben [blank laboratory samples or blank test samples]	69
Leervwert	14, 73
Leistungsparameter, Vergleich	42
line of organic correlation	26
Linearisierung	31
Linearkombinationen	120
LOF	33
log Dosis-Response-Kurve	86
logit-log-Funktion	86
MANDEL	27
Matrixeffekte	112
Matrixnorm	98, 99
maximum likelihood regression	28
Median	28, 29
Mehrfachprobenahme	84
Mehrkomponentensystem	91
Messgrößen	14
Messgrößen Y [responses]	17
Messpräzision	61
Messproben [test portions, analytical portions]	17
Messstellenauswahl	117, 134
Diskriminanzanalyse	136
Faktoranalyse	134
Messstellenbewertung	117
Messwerte [observed values]	14, 18
Methodenauswahl	146
Methodenpräzision	61
Methodenvergleich	25, 38

MHE	84
MLH-Methode	43
Modell der simultanen Gleichungen	116
Modellbewertung	101
Modelle, nichtlineare	117
Modellfehler	41
Modellgleichung	19
Modellgleichungen	101
Multikollinearität	112
Multivariate Kalibration	149
Multivariate Modellierung	149
Nachweisgrenze	61, 65
natürliche Variablen	109
NIPALS-Verfahren	114
NIR-Spektroskopie	112
Normierung auf 100%	71
offset	20
OLS	20
Voraussetzungen	20, 21
Orientierung	86
Methoden	149
Orientierungshilfe	146
Orthogonalmodell	27
P-Matrix-Ansatz	111
P-Matrix-Methode	97
parameterfreie Verfahren	28
PLS-Verfahren, nach WOLD	115
Polar-Diagramme	115
Polarographie	74, 76, 78
Polynomgleichungen	30
Potentiometrie	73, 78, 80, 81
Prädiktorvariable	19
Präzisionsmaße	61
predictive sum of squares	41
PRESS	41
Probe	62, 69
Kalibrierprobe	69
Leerprobe	69
Testprobe	69
Validierprobe	69
Probenahme	17
Probenahmefehler	17
Probensubtraktion	81

Probleme, numerische	113
Problemvermeidung	112
Projektionsmatrix	113
Projektionsmethoden	115
Qualitätssicherung [quality assurance]	
externe	39
interne	38
Rang einer Matrix	113
Rang-Annihilations-Faktoranalyse	115
Rangermittlung	114
Rauschanteil	114
Re-Sampling	40
reduced major axis	26
Referenz, interne	70, 82
Referenzprobe	69
Referenzwerte	40
Regressand	19
Regression	
Alternative Berechnung der Koeffizienten	25
Annahmen	21
Bedingungen	21
'echt nichtlineare'	31
Fehler der Koeffizienten	23
fuzzy	29
gewichtete	21, 28
homogene	23
inverse	26, 72
Modell der reduzierten Hauptachse	26
nichtlineare	21
quasilineare	30
robuste	21, 28, 51
Standardabweichung des Koeffizienten	24
Test der Koeffizienten	24
Test der Signifikanz	32
Vertrauensbereiche der Koeffizienten	25
Regressionsanalyse	149
multiple	148
multivariate (multiple)	149
Regressionsgeraden, Vergleich	59
Regressionskoeffizient, Schätzwert	20
Regressionskoeffizienten	19
Vergleich	42
Regressionsmethoden, Vergleich	42

Regressionsparameter, Vergleich	42
Regressor	19
Rekonstruktionsfehler	118
Reproduzierbarkeit	61
Responsfaktor	84
Restfehler	41
Restliche Zusammensetzung [matrix]	17
Richtigkeit [accuracy]	38, 61
Ridge Regression	113
Ringversuch [cooperative test, round robin]	25, 39
S/R-Modellierung	138
SAM	73
SEC	41
Selektivität	93
SEP	41
SI	16
Signal-Rausch-Optimierung, Kalibrierprobe	119
Signalgrößen [responses]	17
Signalmatrix, Rekonstruktion	117
Signalwert	
kritischer	14, 63
Vertrauensbereich	34, 36
simultanen Gleichungen	116
slope	20
Spektral-Norm	108
Spektralbewertung, visuell	100
Spezifität, Grad der	94
spiking	73
Spurenanalytik	62
SRM	29
Standard	69
Standard-Additionsmethode [standard addition method]	38
Standardaddition	73
Standardmaterial	17
Standardprobe	69
Standardreferenzmaterialien [standard reference material]	38
Standards	17
Standardsubtraktion	80
Statistik	
analytbezogen	58
signalbezogen	56
Sterne-Diagramme	115
Streudiagramme	147

structural relationship	26
Strukturgleichungen	17, 116
Summen der quadratischen Abweichungen [sums of squares of deviation]	20
Systemschätzung	116
test of adequacy	33
Testprobe	69
Testproben-Zugabe	79
Theorie unscharfer Mengen [fuzzy theory]	29
TISAB	74
Transformation	31
Untergrundsubtraktion	112
Ursachenmodellierung	17
Validieren	39
Validierfehler	41
Validierprobe	40, 69
Varianzanalyse [analysis of variances]	34, 56, 149
Varianzanteil	118
Varianzgleichheit	54, 56
Varianzhomogenität	55
Vektornorm	99
Verdünnen	17
Verfahrensstandardabweichung	61
Vergleichspräzision	61
Versuchsplan	92
orthogonaler	133
zweifaktorieller, dreistufiger	134
zweifaktorieller, zweistufiger	133
Versuchsplanung	132
statistische	116
Verteilungsfunktion, Quantil	50
Vertrauensbereich	37, 58
des Signalmittelwerts	37
des Signalwerts	36
eines berechneten Signalwerts	36
Vertrauensbereiche der Vorhersagewerte	36
Vertrauensgrenze [upper confidence limit]	65
Vertrauensintervall	36, 37
Voltammetrie	74
Voraussetzungen, statistische	45
Vorhersagefehler	41
Vorhersagematrix	97
Wahrscheinlichkeit [type I risk] von $\alpha$	66
Wechselwirkungen	116

Wert, kritischer	64
Werte	14
Wertgleichung	20
Wiederholpräzision	61
Wiederholungsmessungen	35
YOUDEN-Plot	39
Zeitreihendarstellungen	147
Zielfunktion	116
Zufällige Verfälschungen	61
Zufallshöchstwert, kritischer	50
Zumischmethode	73
Zusammenhänge von Variablen	148
Zusatzinformation	149
Zuverlässigkeit der Ergebnisse	112
Zweikomponenten-Kalibrierung	89
Zwischen-Klassen-Streumatrix	119

## Danksagung

---

Mit großer Dankbarkeit erfüllt uns die Hilfe derer, die uns in vielfältiger Weise unterstützt haben.

Für die labortechnische Assistenz Liane Aues, die für eine verlässliche Datengrundlage sorgte, ist besonders zu danken. Die Daten, mit denen wir im Mehrkomponenten-Teil arbeiten, beruhen auf ihrer Zuarbeit im Dezember 2019, für deren Erlaubnis Prof. Cepus (Hochschule Merseburg, Fachbereich Ingenieur- und Naturwissenschaften) ebenfalls zu danken ist.

Den Professoren Valentin Cepus und Eckhard Liebscher, deren Lehrgebiete wir mit dem Buch tangieren, danken wir für die kritische Durchsicht des Manuskripts und wertvolle Hinweise.

Da sie beabsichtigen, das Buch in der Lehre zu verwenden, gilt unser vorweggenommener Dank allen studiosi und Kollegen für Hinweise auf übersehene Fehler sowie Hinweise zur Verbesserung.

Der Layouter Yann Labry hat mich (HWZ) endgültig ermutigt, das Buch-Projekt wieder aufzunehmen. Er war sofort in der Lage, den unter WordPerfect Office 2000 geschriebenen Urtext bzw. die daraus erstellte PDF-Datei in die „Layouter-Welt“ zu transferieren und dabei vor allem die zahlreichen Formeln aus dem Quelltext zu übertragen.

Ohne weitere „technische Assistenz“, insbesondere von Stefan Knobloch und Frank Grunicke, wäre das Buch so schnell nicht fertig geworden.

Am Ende aber haben wir es wohl Dr. Frank Baumann zu verdanken, dass das Buch entstanden ist. Er erinnerte mich (HWZ) immer wieder an mein Versprechen, mit einer Publikation zur Analytik das Angebotsspektrum des Hochschulverlags zu bereichern. (Allerdings gab ich ihm das Versprechen vor meiner schweren Erkrankung...)

Schließlich gilt unser Dank der Beschaffungsstelle der Hochschule, die unsere Wünsche meist zeitnah erfüllen konnte.