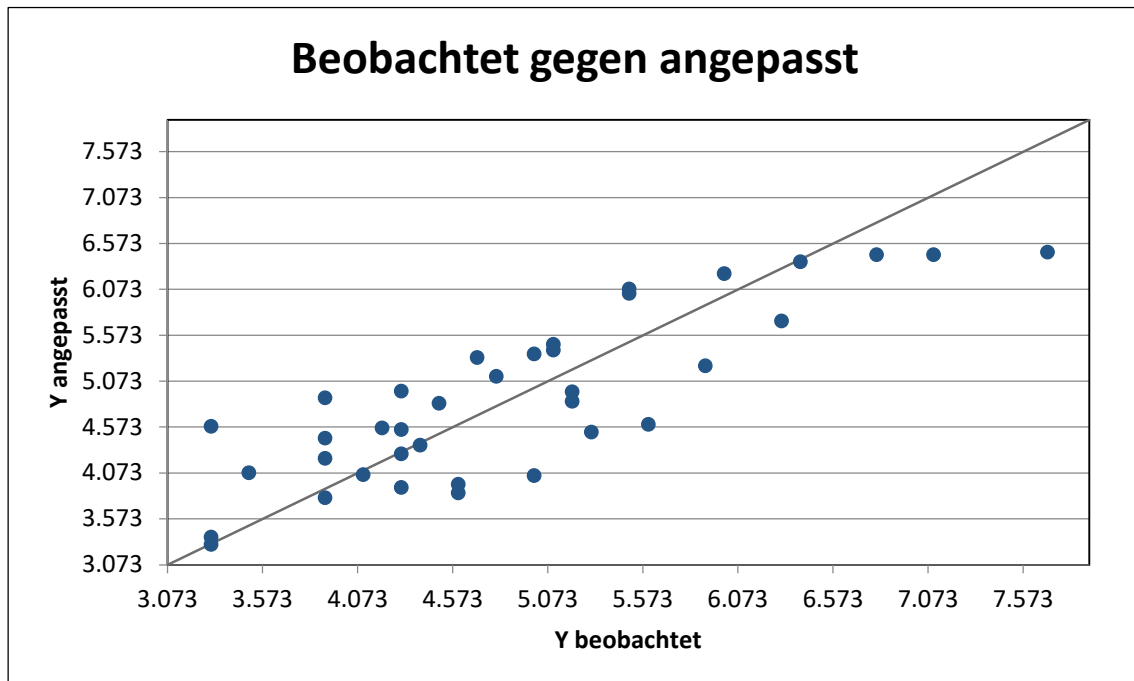


Partial Least Squares-Regression

Methodik, Modellvalidierung und Interpretation



Eckehardt Spenhoff

Alle Rechte, auch die Übersetzung in fremden Sprachen, vorbehalten. Kein Teil dieses Werkes darf ohne vorherige schriftliche Genehmigung des Verfassers in irgendeiner Form, auch nicht zum

Zweck der Unterrichtsgestaltung, reproduziert oder unter Verwendung elektronischer Medien verarbeitet und vervielfältigt werden.

Copyright © 2026 Eckehardt Spenhoff

Inhaltsverzeichnis

Anwendungsbeispiel zur Software OQM-Stat.....	5
1. Problemstellung und Ziel der PLS1-Regression.....	5
2. Datenbasis und sachliche Einordnung.....	5
3. Datenvorbereitung und Standardisierung.....	6
4. Komponentenbildung nach dem NIPALS-Prinzip.....	6
5. Rücktransformation und Regressionsgleichung.....	7
6. Modellanpassung: SSE, SST und R^2	7
7. Kreuzvalidierung und Prognoseleistung.....	8
8. Die Modellauswahltabelle im Aroma-Beispiel.....	8
9. Korrigiertes R^2 , RMSE, AIC und BIC.....	9
10. Varianzanalyse und globaler Modelltest.....	9
11. Koeffizienten und ihre inhaltliche Bedeutung.....	10
12. Jackknife-Standardfehler und Konfidenzintervalle.....	10
13. Variable Importance in Projection (VIP).....	11
14. Residuen und Vorhersageresiduen.....	11
15. Hebelwirkung, X-Distanz und X-Residualdistanz.....	12
16. Interpretation der vier Diagnosegrafiken.....	12
17. Einordnung der Aroma-Ergebnisse.....	13
18. Reproduzierbarkeit in OQM-Stat.....	13
19. Grenzen und sinnvolle Erweiterungen.....	14
20. Schlussfolgerung.....	14

Anwendungsbeispiel zur Software OQM-Stat

Die Partial Least Squares Regression mit einer Zielgröße (PLS1) verdichtet korrelierte Prädiktoren zu latenten Komponenten, die gezielt auf den Zusammenhang mit der Zielgröße ausgerichtet sind. Der Artikel entwickelt die Rechenschritte der in OQM-Stat implementierten NIPALS-artigen PLS1-Regression und erläutert Kreuzvalidierung, Modellauswahl, Regressionskoeffizienten, Jackknife, VIP und Beobachtungsdiagnostik. Die vollständige Ergebnisinterpretation orientiert sich an einem Datensatz mit 37 Aroma-Bewertungen und 17 Elementkonzentrationen. Das ausgewählte Zweikomponentenmodell erklärt 70,14 % der Aroma-Variation; das kreuzvalidierte prognostische Bestimmtheitsmaß beträgt 46,32 %.

1. Problemstellung und Ziel der PLS1-Regression

Bei multivariaten Messungen enthalten die Prädiktoren häufig gemeinsame Information. Besonders chemische Konzentrationsprofile können stark korrelieren, weil sie auf gemeinsame Herkunft, Produktionsbedingungen oder analytische Zusammenhänge zurückgehen. In der gewöhnlichen multiplen Regression kann dies die Einzelkoeffizienten instabil machen. PLS1 konstruiert stattdessen wenige latente Komponenten aus allen Prädiktoren und passt die Zielgröße an diese Komponenten an.

Die Bezeichnung PLS1 bedeutet, dass genau eine abhängige Variable Y modelliert wird. Mehrere Zielgrößen würden eine andere Modellvariante erfordern. Die Zahl der Ausgangsprädiktoren p bleibt von der Zahl a der tatsächlich genutzten Komponenten zu unterscheiden. Ein Modell mit $a = 2$ nutzt weiterhin die Information sämtlicher 17 Elementmessungen; es setzt nicht etwa 15 Regressionskoeffizienten auf null.

Ein wesentlicher Unterschied zur Hauptkomponentenregression (PCR) besteht darin, dass PCR Richtungen maximaler X-Streuung ohne Berücksichtigung von Y sucht. PLS1 nutzt die Kovariation mit Y bereits beim Aufbau der Komponenten. Deshalb kann ein kleiner Anteil erklärter X-Varianz mit einem deutlich größeren Anteil erklärter Y-Varianz einhergehen.

2. Datenbasis und sachliche Einordnung

Die Untersuchung umfasst 37 Beobachtungen und 17 Prädiktoren: Cd, Mo, Mn, Ni, Cu, Al, Ba, Cr, Sr, Pb, B, Mg, Si, Na, Ca, P und K. Die Zielvariable trägt die Bezeichnung „Aroma“. Es wird angenommen, dass die Werte aus einer sensorischen Weinbeurteilung stammen. Eine Dokumentation des Bewertungsverfahrens, der Prüferzahl und der Einheit der chemischen Messungen liegt nicht vor. Aus diesem Grund werden die Elementwerte als Messgrößen ohne behauptete Einheit und Aroma als numerische Bewertung interpretiert.

Die beobachteten Aroma-Werte reichen von 3,3 bis 7,7. Die Prädiktoren besitzen sehr unterschiedliche Größenordnungen, beispielsweise liegen einige Spurenelementwerte unter 0,1, während K im Bereich mehrerer Hundert bis über Tausend liegt. Eine Standardisierung ist daher wesentlich, damit die Komponenten nicht allein von großen numerischen Skalen dominiert werden.

Merkmal	Aroma-Fallstudie	Bedeutung
Beobachtungen n	37	Unabhängige Datensätze
Prädiktoren p	17	Elementkonzentrationen
X-Rang	17	Volle numerische Spaltenrangzahl
CV-Gruppengröße	1	Leave-one-out
Gewählte Komponenten	2	Minimum von PRESS

Zielgröße	Aroma	Numerische Bewertung
-----------	-------	----------------------

3. Datenvorbereitung und Standardisierung

OQM-Stat zentriert jeden Prädiktor an seinem Stichprobenmittel und dividiert durch seine Stichprobenstandardabweichung mit dem Nenner $n - 1$. Die Zielgröße wird dagegen nur zentriert. Damit bleibt ihre natürliche Bewertungsskala für die spätere Rücktransformation erhalten. Standardisierung der X-Variablen:

$$x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{s_{x_j}}, \quad s_{x_j} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$$

Hinweis: Jede Spalte erhält Mittelwert null und Stichprobenstandardabweichung eins.

Standardisierung der X-Variablen. Jede Spalte erhält Mittelwert null und Stichprobenstandardabweichung eins. Zentrierung der Zielgröße:

$$y_i^* = y_i - \bar{y}$$

Hinweis: Y wird nicht durch seine Standardabweichung dividiert.

Konstante Prädiktoren sind für die standardisierte Komponentenbildung ungeeignet. Die Kernroutine beendet die Berechnung, falls die zentrierte Quadratsumme einer X-Spalte numerisch zu klein ist. In der Aroma-Analyse tritt dieser Fall nicht auf.

4. Komponentenbildung nach dem NIPALS-Prinzip

Die Implementierung verwendet ein sequenzielles, auf eine Zielvariable zugeschnittenes PLS1-Verfahren. Aus der jeweils aktuellen X-Residualmatrix und dem aktuellen Y-Residualvektor wird zunächst ein Gewichtsvektor bestimmt. Anschließend entstehen die Scores als lineare Kombinationen der Prädiktoren. Die X- und Y-Ladungen werden durch Projektion auf diese Scores berechnet. Berechnung des Gewichtsvektors:

$$\mathbf{w}_a = \frac{\mathbf{X}_{a-1}^T \mathbf{y}_{a-1}}{\|\mathbf{X}_{a-1}^T \mathbf{y}_{a-1}\|_2}$$

Hinweis: Der Gewichtsvektor hat euklidische Länge eins.

Berechnung des Scorevektor:

$$\mathbf{t}_a = \mathbf{X}_{a-1} \mathbf{w}_a$$

Hinweis: Die Scorewerte beschreiben die Positionen der Beobachtungen auf Komponente a.

Berechnung der X- und Y-Ladungen:

$$\mathbf{p}_a = \frac{\mathbf{X}_{a-1}^T \mathbf{t}_a}{\mathbf{t}_a^T \mathbf{t}_a}, \quad q_a = \frac{\mathbf{y}_{a-1}^T \mathbf{t}_a}{\mathbf{t}_a^T \mathbf{t}_a}$$

Hinweis: Die Ladungen charakterisieren die Beziehung der Komponente zu X und Y.

Berechnung der Deflation:

$$\mathbf{X}_a = \mathbf{X}_{a-1} - \mathbf{t}_a \mathbf{p}_a^T, \quad \mathbf{y}_a = \mathbf{y}_{a-1} - q_a \mathbf{t}_a$$

Hinweis: Nach jeder Komponente wird der bereits erklärte Anteil aus X und Y entfernt.

Die nächste Komponente wird aus den verbleibenden Residuen gebildet. Dieses schrittweise Verfahren bewirkt, dass aufeinanderfolgende Scorevektoren orthogonal sind, soweit die numerische Genauigkeit dies zulässt. Es erklärt zugleich, weshalb die kumulative X-Varianz mit jeder Komponente steigt.

5. Rücktransformation und Regressionsgleichung

Die Scores sind Hilfsgrößen für die Schätzung. Für die praktische Anwendung benötigt man eine Gleichung in den ursprünglichen Elementkonzentrationen. OQM-Stat berechnet aus Gewichten, X-Ladungen und Y-Ladungen zunächst die Koeffizienten für standardisierte X-Werte. Anschließend werden die Koeffizienten mit den ursprünglichen X-Standardabweichungen zurücktransformiert und der Achsenabschnitt bestimmt. Berechnung der PLS-Koeffizienten:

$$\beta^* = W(P^T W)^{-1} q$$

Hinweis: Die im Code verwendete Dreieckslösung entspricht dieser Matrixdarstellung.

Berechnung der Originalskala und Konstante:

$$\beta_j = \frac{\beta_j^*}{s_{x_j}}, \quad \beta_0 = \bar{y} - \sum_{j=1}^p \beta_j \bar{x}_j$$

Hinweis: Die ausgegebene Regressionsgleichung verwendet die ursprünglichen X-Werte.

Berechnung der Vorhersagewerte:

$$\hat{y}_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij}$$

Hinweis: Die Koeffizienten aller 17 Elemente gehen in die Aroma-Vorhersage ein.

Die zusätzlich ausgegebenen standardisierten Koeffizienten setzen den Effekt auf eine Änderung um eine Standardabweichung von X und messen die Veränderung in Standardabweichungen von Y. Sie erlauben einen Skalenvergleich zwischen den Elementen, sind jedoch nicht mit VIP identisch.

6. Modellanpassung: SSE, SST und R²

Die Anpassungsgüte beschreibt, wie eng die berechneten Aroma-Werte bei den beobachteten liegen. OQM-Stat bestimmt für jede Komponentenanzahl die Fehlerquadratsumme SSE und vergleicht sie mit der gesamten zentrierten Quadratsumme SST. Berechnung der Quadratsummen:

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad SST = \sum_{i=1}^n (y_i - \bar{y})^2$$

Hinweis: SSE misst die verbleibende Abweichung, SST die Gesamtstreuung.

Berechnung des Bestimmtheitsmaß:

$$R^2 = 1 - \frac{SSE}{SST}$$

Hinweis: R² steigt im Allgemeinen, wenn weitere Komponenten hinzukommen.

Im Aroma-Beispiel beträgt $SST = 41,203243$. Mit zwei Komponenten ergibt sich $SSE = 12,304356$ und damit $R^2 = 0,701374$. Die beiden Komponenten erklären rund 70,1 % der beobachteten Aroma-Streuung. Die Anpassung ist deutlich besser als die alleinige Verwendung des Stichprobenmittelwertes.

7. Kreuzvalidierung und Prognoseleistung

Die zentrale Frage der Modellwahl lautet nicht, mit wie vielen Komponenten sich die vorhandenen Beobachtungen am besten anpassen lassen, sondern wie gut das Modell ausgelassene Beobachtungen vorhersagt. OQM-Stat verwendet bei CV-Gruppengröße eins die Leave-one-out-Kreuzvalidierung (LOOCV). Für jede Beobachtung wird das Modell auf den übrigen 36 Beobachtungen neu berechnet, einschließlich Mittelwertbildung, Standardisierung und PLS-Komponentenbildung.

Für jede mögliche Komponentenanzahl entsteht so ein vollständiger Satz prognostischer Werte. Ihre Abweichungen von den beobachteten Aroma-Werten ergeben die PRESS-Summe. Wichtig ist, dass die ausgelassene Beobachtung nicht an der Berechnung ihrer eigenen Vorhersage beteiligt ist. Berechnung des PRESS:

$$PRESS(a) = \sum_{i=1}^n \left(y_i - \hat{y}_{i,-i}^{(a)} \right)^2$$

Hinweis: a bezeichnet die Komponentenanzahl, -i das ohne Beobachtung i geschätzte Modell.

Berechnung des Prognostischen R^2 :

$$R_{\text{prog}}^2(a) = 1 - \frac{PRESS(a)}{SST}$$

Hinweis: Der Code verwendet die Gesamt-SST als Referenzgröße.

Die automatische Wahl erfolgt strikt über das kleinste PRESS. Anders als bei einer One-standard-error-Regel wird keine zusätzliche Einfachheitsregel angewendet. Wenn zwei PRESS-Werte exakt gleich sind, bleibt aufgrund des strikten Kleiner-als-Vergleichs die zuerst gefundene, also kleinere Komponentenanzahl ausgewählt.

8. Die Modellauswahltabelle im Aroma-Beispiel

Komponenten	X-Varianz	SSE	R^2	PRESS	R^2 prog.
1	0.225149	16.540251	0.598569	22.390448	0.456585
2	0.366697	12.304356	0.701374	22.116331	0.463238
3	0.437655	8.93848	0.781720	23.305510	0.434371
4	0.537093	8.276141	0.799139	22.261044	0.459726

Der Übergang von einer auf zwei Komponenten reduziert die Anpassungsfehlerquadratsumme von 16,540 auf 12,304 und verbessert zugleich PRESS leicht von 22,390 auf 22,116. Eine dritte Komponente verbessert zwar das gewöhnliche R^2 von 0,701 auf 0,782, verschlechtert aber PRESS auf 23,306. Die dritte Komponente bildet also zusätzliche Eigenschaften der vorhandenen Stichprobe ab, ohne deren LOOCV-Vorhersage zu verbessern.

Die PRESS-Differenz zwischen einer und zwei Komponenten beträgt nur etwa 0,274. Der prognostische Zugewinn ist daher moderat. Gleichwohl folgt die automatische Auswahlregel eindeutig dem Minimum bei zwei Komponenten. Das Beispiel zeigt, warum ein steigendes R^2 allein kein geeignetes Kriterium für die Wahl der PLS-Komplexität ist. Interessant ist, dass der PRESS bei 4 Kompo-

nenten nur geringfügig höher ist, bei gleichzeitiger Erhöhung des Bestimmtheitsmaßes. Möchte man diese Modell analysieren, dann muss die Anzahl der Komponenten eingegeben werden. Die kumulative X-Varianz der beiden Komponenten beträgt 36,67 %, während das Modell 70,14 % der Aroma-Variation erklärt. PLS sucht somit gezielt jene Kombinationen chemischer Merkmale, die für die Aroma-Bewertung relevant sind, statt möglichst viel Gesamtstreuung der chemischen Messmatrix zu erfassen.

9. Korrigiertes R^2 , RMSE, AIC und BIC

OQM-Stat ergänzt die Anpassungs- und Prognosemaße um Kennzahlen, die eine Komplexitätskorrektur enthalten. Dabei wird die Komponentenanzahl a als Näherung für die Modellfreiheitsgrade verwendet. Die Kennzahlen sind in der Software entsprechend mit einem Stern gekennzeichnet.

Berechnung des Korrigierten R^2 :

$$R_{\text{kor}}^2 = 1 - \frac{\text{SSE} / (n - a - 1)}{\text{SST} / (n - 1)}$$

Hinweis: Mit $a = 2$ und $n = 37$ ergibt sich 0,683808.

Berechnung des RMSE der Anpassung:

$$\text{RMSE} = \sqrt{\frac{\text{SSE}}{n - a - 1}}$$

Hinweis: Der Wert 0,601575 liegt in Einheiten der Aroma-Bewertung.

Berechnung von AIC* und BIC*:

$$\text{AIC}^* = n \ln(\text{SSE} / n) + 2k, \quad \text{BIC}^* = n \ln(\text{SSE} / n) + k \ln n, \quad k = a + 2$$

Hinweis: Die gemeinsame Normalverteilungskonstante wird ausgelassen; k ist approximativ.

Kennzahl	Zweikomp.-modell	Interpretation
R^2	0,701374	Erklärte Variation in der Stichprobe
R^2 korrigiert*	0,683808	Anpassung mit Freiheitsgradkorrektur
R^2 prognostisch	0,463238	LOOCV-Prognose im Vergleich zur Gesamtmittelwert-Basis
RMSE*	0,601575	Typische Reststreuung auf Aroma-Skala
AIC*	-32,735688	Approximatives Informationskriterium
BIC*	-26,292016	Stärkerer Komplexitätszuschlag

AIC* und BIC* werden in der Tabelle für mehrere Komponenten ausgewiesen und können die Anpassungsentwicklung ergänzend charakterisieren. Sie ersetzen in der hier implementierten automatischen Auswahl nicht PRESS. Die auf $a + 2$ beruhende Parameterzählung ist eine pragmatische Näherung, weil die effektive Flexibilität einer PLS-Schätzung von der datenabhängigen Komponentenbildung beeinflusst wird.

10. Varianzanalyse und globaler Modelltest

Die Varianzanalyse zerlegt die Gesamtquadratsumme in erklärte und unerklärte Variation. Für das ausgewählte Aroma-Modell verwendet OQM-Stat $a = 2$ Regressionsfreiheitsgrade und $n - a - 1 = 34$ Residualfreiheitsgrade. Daraus ergibt sich ein F-Verhältnis von rund 39,93.

Berechnung des F-Verhältnis.

$$F = \frac{(SST - SSE) / a}{SSE / (n - a - 1)}$$

Hinweis: Die Freiheitsgrade sind für PLS als Approximation gekennzeichnet.

Quelle	DF	SS	MS	F
Regression	2	28,898887	14,449443	39,927410
Residuen	34	12,304356	0,361893	
Gesamt	36	41,203243		

Das große F-Verhältnis spiegelt den deutlichen Anpassungsgewinn gegenüber einem konstanten Modell wider. Der in der Ausgabe auf sechs Dezimalstellen als 0,000000 dargestellte p-Wert bedeutet nicht $p = 0$, sondern einen sehr kleinen Wert. Weil die PLS-Komponenten aus X und Y geschätzt werden, sind die konventionellen F-Verteilungsannahmen nicht unmittelbar mit denen einer Regression auf vorab festgelegte Prädiktoren gleichzusetzen. OQM-Stat bezeichnet den Test daher ausdrücklich als approximativ.

11. Koeffizienten und ihre inhaltliche Bedeutung

Die unstandardisierten Regressionskoeffizienten beschreiben die Änderung der geschätzten Aroma-Bewertung bei einer Erhöhung des jeweiligen Elements um eine Messeinheit unter Konstanzhaltung der anderen Elemente innerhalb der linearen Modellgleichung. Wegen der starken gemeinsamen Struktur der chemischen Variablen sind solche bedingten Einzelwirkungen nicht mit isolierten chemischen Ursache-Wirkungs-Effekten gleichzusetzen.

Die standardisierten Koeffizienten erlauben einen unmittelbaren Größenvergleich. Für Sr beträgt der standardisierte Koeffizient etwa $-0,316$; für B etwa $+0,226$ und für Mo etwa $+0,164$. Das Vorzeichen beschreibt die Richtung des bedingten Modellzusammenhangs, nicht die Gesamtassoziation ohne Berücksichtigung anderer Elemente.

Element	Koeffizient	standardisiert	p (JK)	VIP
Sr	-0,764497	-0,315640	0,000195	2,111154
Ba	-1,694845	-0,165764	0,010266	1,558483
Mg	-0,005021	-0,139192	0,001562	1,416724
K	-0,000521	-0,092910	0,207127	1,237302
B	+0,145598	+0,225782	0,042079	1,153380
Ca	-0,007867	-0,141744	0,046701	1,104426
Na	-0,000824	-0,023331	0,811585	1,007103
Cd	-0,882959	-0,020243	0,824499	0,989057
Mo	+1,797245	+0,164133	0,028940	0,818031
Pb	+0,306505	+0,134004	0,037236	0,670496

Die Tabelle illustriert, dass Rangfolgen nach standardisiertem Koeffizienten, Jackknife-p-Wert und VIP nicht übereinstimmen müssen. Sr weist sowohl einen großen negativen standardisierten Koeffizienten als auch den höchsten VIP auf. Mo besitzt einen positiven, nach Jackknife auffälligen Koeffizienten.

fizienten, erreicht aber keinen VIP über eins. K hat einen VIP über eins, während sein Jackknife-p-Wert deutlich über 0,05 liegt. Diese Befunde beantworten unterschiedliche Fragen.

12. Jackknife-Standardfehler und Konfidenzintervalle

OQM-Stat führt für jeden Koeffizienten $n = 37$ Leave-one-out-Neuschätzungen mit der zuvor gewählten Komponentenanzahl durch. Dabei wird die Komponentenanzahl nicht in jeder Teilstichprobe erneut optimiert. Aus der Streuung dieser Schätzungen entstehen die Jackknife-Standardfehler. Die ausgegebenen t-Werte, p-Werte und 95%-Intervalle sind entsprechend als approximativ bezeichnet. Berechnen des Jackknife-Standardfehler:

$$SE_{JK}(\hat{\beta}_j) = \sqrt{\frac{n-1}{n} \sum_{i=1}^n (\hat{\beta}_{j,(-i)} - \bar{\hat{\beta}}_{j,(-)})^2}$$

Hinweis: Die Formel entspricht der implementierten Jackknife-Skalierung.

Berechnung von Jackknife-t und Konfidenzintervall:

$$t_j = \frac{\hat{\beta}_j}{SE_{JK}(\hat{\beta}_j)}, \quad \hat{\beta}_j \pm t_{0,975,n-1} SE_{JK}(\hat{\beta}_j)$$

Hinweis: OQM-Stat verwendet $n - 1$ Freiheitsgrade.

Beispielsweise beträgt das 95%-Jackknife-Intervall für Sr etwa $[-1,138188; -0,390807]$ und für Ba $[-2,962698; -0,425991]$. Beide Intervalle liegen vollständig unter null. Für Mo ergibt sich dagegen ein positives Intervall $[0,195288; 3,399202]$. Die Angaben sind eine hilfreiche Beschreibung der Koeffizientenstabilität bei Weglassen einzelner Beobachtungen.

13. Variable Importance in Projection (VIP)

Der VIP-Wert fasst zusammen, wie stark ein Prädiktor über sämtliche berücksichtigten Komponenten hinweg an der Erklärung der Zielgröße beteiligt ist. OQM-Stat gewichtet dazu die quadrierten Gewichte der Komponenten mit deren Beitrag zur erklärten Y-Quadratsumme. Anschließend wird über die Prädiktorzahl p normiert. Berechnung der VIP:

$$VIP_j = \sqrt{p \frac{\sum_{a=1}^A SSY_a w_{ja}^2}{\sum_{a=1}^A SSY_a}}, \quad SSY_a = q_a^2 \mathbf{t}_a^T \mathbf{t}_a$$

Hinweis: Die Gewichte w sind in jeder Komponente auf Länge eins normiert.

VIP-Werte über eins gelten häufig als Hinweis auf eine überdurchschnittliche Beteiligung an der PLS-Modellierung von Y. Diese Schwelle ist eine Interpretationskonvention und kein Signifikanztest. Für die Aroma-Daten ist Sr mit $VIP = 2,111$ am wichtigsten, gefolgt von Ba (1,558), Mg (1,417), K (1,237), B (1,153), Ca (1,104) und Na (1,007).

Der Vergleich von Na und Sr macht die Aussagekraft deutlich: Beide überschreiten die VIP-Schwelle eins, doch Na besitzt einen sehr kleinen standardisierten Koeffizienten und einen hohen Jackknife-p-Wert. VIP misst einen Beitrag zum Komponentenmodell, nicht die Sicherheit eines isolierten bedingten Regressionskoeffizienten. Daher ist eine gemeinsame Betrachtung von VIP, Koeffizienten und Intervallen sachgerecht.

14. Residuen und Vorhersageresiduen

Die Beobachtungsdiagnostik unterscheidet zwischen Anpassungsresiduen und prognostischen Residuen. Das Anpassungsresiduum vergleicht den beobachteten Wert mit einer Schätzung aus dem Modell, in dessen Anpassung die Beobachtung enthalten war. Das prognostische Residuum verwendet dagegen die aus der Kreuzvalidierung stammende Vorhersage. Berechnung der Residuen:

$$e_i = y_i - \hat{y}_i, \quad e_{i,\text{prog}} = y_i - \hat{y}_{i,-i}$$

Hinweis: Die Quadratsumme der prognostischen Residuen ergibt PRESS.

Berechnung des Standardisierten Residuum:

$$r_i = \frac{e_i}{\text{RMSE} \sqrt{1 - h_i}}$$

Hinweis: Die Implementierung berücksichtigt die ausgegebene Hebelwirkung h .

Ein Residuum nahe null bedeutet eine gute Übereinstimmung zwischen Beobachtung und Anpassung. Positive Residuen kennzeichnen eine Unterschätzung des Aromas, negative eine Überschätzung. Große prognostische Residuen zeigen, bei welchen Proben die Vorhersage ohne Kenntnis dieser Probe weniger gut gelingt.

15. Hebelwirkung, X-Distanz und X-Residualdistanz

Die PLS-Scores beschreiben die Position einer Probe im latenten Komponentenraum. OQM-Stat berechnet daraus die Hebelwirkung sowie eine skalierte X-Distanz. Zusätzlich wird die nach Deflation verbleibende X-Residualquadratsumme pro Beobachtung ausgegeben. Damit lassen sich ungewöhnliche Proben nach ihrer Lage im Modellraum und nach ihrer durch die Komponenten nicht erfassten X-Struktur unterscheiden. Berechnung der Hebelwirkung:

$$h_i = \frac{1}{n} + \sum_{a=1}^A \frac{t_{ia}^2}{\mathbf{t}_a^T \mathbf{t}_a}$$

Hinweis: Die Hebelwirkung enthält den Achsenabschnittsbeitrag $1/n$.

Formel (24) –X-Distanz und X-Residualdistanz.

$$D_{X,i} = (n-1) \sum_{a=1}^A \frac{t_{ia}^2}{\mathbf{t}_a^T \mathbf{t}_a}, \quad Q_{X,i} = \sum_{j=1}^p (x_{ij,A}^*)^2$$

Hinweis: Q ist die Summe der quadrierten standardisierten X-Residuen nach A Deflationsschritten.

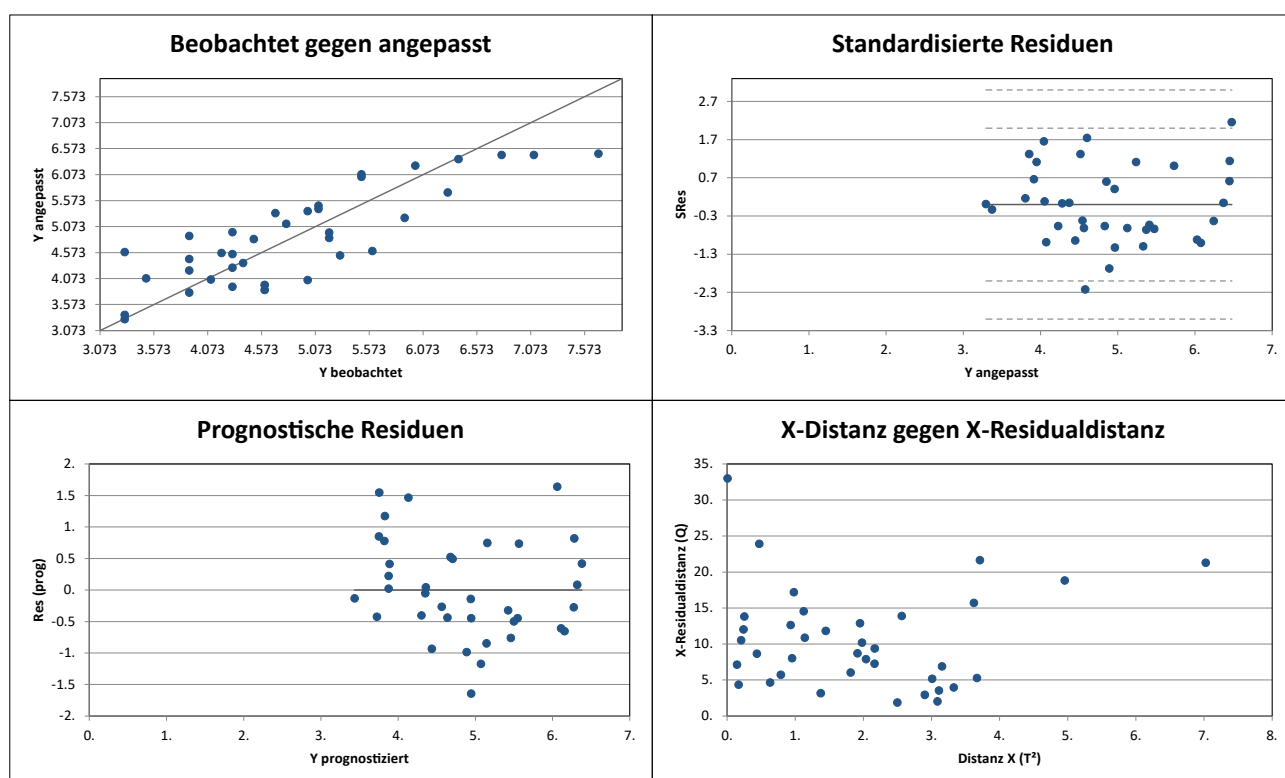
Eine hohe X-Distanz bedeutet eine weit vom Zentrum entfernte Position im durch die PLS-Komponenten beschriebenen Raum. Eine hohe X-Residualdistanz bedeutet, dass das chemische Profil einer Probe durch diese Komponenten nur unvollständig beschrieben wird. Beide Eigenschaften können bei derselben Probe auftreten, müssen es aber nicht.

16. Interpretation der vier Diagnosegrafiken

Grafik	Achsen / Darstellung	Worauf achten?
Beobachtet gegen angepasst	Aroma gegen Modellanpassung	Nähe zur Diagonalen; systematische Abweichungen
Standardisierte Residuen gegen	Residuen in RMSE-Einheiten	Muster, Streuungsänderungen

Anpassung	gegen Schätzwert	und auffällige Beobachtungen
Prognostische Residuen gegen Prognose	LOOCV-Fehler gegen LOOCV-Schätzwert	Vorhersagefehler und mögliche systematische Verzerrungen
X-Distanz gegen X-Residualdistanz	Lage im Komponentenraum gegen nicht erklärte X-Struktur	Proben mit außergewöhnlichen chemischen Profilen

Die erste Grafik visualisiert unmittelbar die Anpassung des Modells. Eine ideale Anpassung würde alle Punkte auf die Winkelhalbierende legen. Die zweite Grafik sollte im Idealfall eine unstrukturierte Streuung um die Nulllinie zeigen. Die dritte Grafik ist für die Vorhersagebeurteilung besonders wichtig, weil jede Prognose aus einem Modell ohne die jeweils betrachtete Probe stammt. Die vierte Grafik ergänzt die Y-bezogenen Diagnosen um eine Sicht auf die chemische X-Struktur. Die vier Diagramme erfüllen unterschiedliche Aufgaben und sollten nicht durch einen einzigen globalen Gütewert ersetzt werden.



17. Einordnung der Aroma-Ergebnisse

Das Zweikomponentenmodell stellt einen bemerkenswert kompakten Zusammenhang zwischen 17 Elementkonzentrationen und der Aroma-Bewertung dar. Mit $R^2 = 0,701$ gelingt eine gute Anpassung innerhalb der vorhandenen Stichprobe. Das niedrigere prognostische $R^2 = 0,463$ zeigt, dass die Leave-one-out-Vorhersagen erwartungsgemäß anspruchsvoller sind als die Anpassung derselben Daten. Das Ergebnis ist daher als nützlicher, aber nicht vollständig ausgeschöpfter prognostischer Zusammenhang zu lesen.

Unter den chemischen Merkmalen hebt sich Sr besonders deutlich hervor. Sein negativer standardisierter Koeffizient, sein hoher VIP und das negative Jackknife-Konfidenzintervall weisen übereinstimmend auf eine wichtige Rolle im gewählten Modell hin. Auch Ba und Mg zeigen ausgeprägte negative Beiträge. B und Mo weisen positive bedingte Koeffizienten auf. Die Beziehungen sind Eigenschaften der vorliegenden Stichprobe und der gewählten linearen Modellstruktur.

Die Daten erlauben keine Aussage darüber, ob einzelne Elemente unmittelbar das sensorische Aroma verursachen. Möglich ist ebenso, dass Elementkonzentrationen Herkunfts- oder Prozessmerkmale abbilden, die mit der Aroma-Bewertung zusammenhängen. Für die statistische Zielsetzung ist entscheidend, dass die chemischen Profile Information enthalten, die zur Schätzung des Aromas genutzt werden kann.

18. Reproduzierbarkeit in OQM-Stat

Die Reihenfolge der Software lautet: numerische Eingabebereiche prüfen; Rang der X-Matrix bestimmen; PLS-Modelle für zulässige Komponentenanzahlen schätzen; Leave-one-out-PRESS berechnen; optimale Komponentenanzahl auswählen; vollständiges Modell neu schätzen; Kennzahlen und approximative ANOVA berechnen; Jackknife durchführen; VIP und Beobachtungsdiagnostik ausgeben; Diagramme erzeugen.

Bei CV-Gruppengröße eins ist die Zuordnung der ausgelassenen Beobachtungen deterministisch. Für größere Gruppen mischt der Code die Beobachtungsindizes mit einem Fisher-Yates-Verfahren. In solchen Fällen kann die konkrete Gruppeneinteilung das CV-Ergebnis beeinflussen. Im vorliegenden Aroma-Beispiel mit Gruppengröße eins besteht diese Zufallsabhängigkeit nicht.

Die Software überprüft zusätzlich, ob die Summe der quadrierten prognostischen Residuen mit dem zuvor berechneten PRESS übereinstimmt. Dieser interne Abgleich stellt eine nützliche Konsistenzprüfung zwischen Modellübersicht und Beobachtungstabelle dar.

19. Grenzen und sinnvolle Erweiterungen

Die wichtigste fachliche Grenze des Beispiels ist die fehlende Dokumentation der sensorischen Erhebung. Ohne Angaben zur Verkostungsmethode lässt sich die Bedeutung eines Aroma-Unterschieds von beispielsweise 0,6 Punkten nicht unmittelbar in eine sensorische Qualitätsstufe übersetzen. Auch die Messeinheiten der Elemente sind nicht angegeben. Die statistische Auswertung selbst bleibt davon unberührt, die fachliche Interpretation der Regressionskoeffizienten aber bewusst zurückhaltend.

Eine externe Validierungsstichprobe wäre eine sinnvolle Ergänzung, wenn das Modell zur Vorhersage neuer Weinproben eingesetzt werden soll. Ebenso könnte eine Wiederholungsmessung der Aroma-Bewertungen die sensorische Messunsicherheit quantifizieren. Beides ist eine Perspektive für weitere Untersuchungen und keine Voraussetzung für die Aussagekraft der hier vorliegenden explorativen PLS1-Analyse.

20. Schlussfolgerung

Die PLS1-Regression von OQM-Stat verbindet eine nachvollziehbare latente Komponentenbildung mit einer konsequenten Kreuzvalidierung und differenzierten Diagnosemöglichkeiten. Das Aroma-Beispiel zeigt den praktischen Nutzen: Aus 17 Elementmessungen entstehen zwei Komponenten, die den größten Teil der innerhalb der Stichprobe erkläraren Variation erfassen und zugleich eine messbare Prognoseleistung besitzen. Besonders aufschlussreich ist die gemeinsame Interpretation von R^2 , PRESS, prognostischem R^2 , Koeffizienten, Jackknife-Intervallen und VIP-Werten. Erst ihr Zusammenspiel liefert ein tragfähiges Bild des Modells.

Literaturhinweise

- Wold, H. (1966): Estimation of principal components and related models by iterative least squares.
- Geladi, P. und Kowalski, B. R. (1986): Partial least-squares regression: a tutorial. *Analytica Chimica Acta* 185, 1–17.
- Martens, H. und Næs, T. (1989): *Multivariate Calibration*. Wiley.
- Mevik, B.-H. und Wehrens, R. (2007): The pls Package: Principal Component and Partial Least Squares Regression in R. *Journal of Statistical Software* 18(2).