

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/46751>

Please be advised that this information was generated on 2020-10-21 and may be subject to change.

Einführung in die Funktionsweise des PLS-Algorithmus

Jörg Betzin¹ und Jörg Henseler²

1	Einführung	50
1.1	Notation	51
1.2	Die Modellgleichungen im Pfadmodell	52
2	Der PLS-Ansatz zur Schätzung der latenten Variablen	54
2.1	Der ALS Algorithmus zur Bestimmung der ersten Hauptkomponente . . .	56
2.2	Der ALS-Algorithmus in der Kanonischen Korrelationsanalyse	58
2.3	Der PLS Schätzalgorithmus	60
3	Die Lösung des Strukturgleichungsmodells	69

¹Dr. Jörg Betzin ist wissenschaftlicher Assistent am Institut für Quantitative Methoden der Technischen Universität Berlin.

²Dipl.-Wirtsch.-Ing. Jörg Henseler ist Assistant Professor für Marketing an der Nijmegen School of Management, Radboud Universiteit Nijmegen, Niederlande.

1 Einführung

Die Anwendung von Strukturgleichungsmodellen dient der Analyse komplexer Wirkzusammenhänge zwischen Konstrukten (latenten Variablen – *LV*), die i. d. R. mit Hilfe mehrerer Indikatoren (manifesten Variablen – *MV*) gemessen werden. Diese Wirkzusammenhänge werden als kausale Relationen angenommen und können graphisch in einem Pfadmodell dargestellt werden. Abbildung 1.1 zeigt ein Beispiel für ein Pfadmodell.

Die allgemeine Form von Strukturgleichungsmodellen ist die Modellierung der kausalen Zusammenhänge sowohl der *LV* untereinander als auch zwischen den *LV* und *MV* in einem gemeinsamen System von Struktur- und Messgleichungen. Die Struktur- und Messgleichungen werden dabei formal in zwei Teilmodellen angegeben. Das Strukturgleichungssystem (*SGS*) beschreibt den Zusammenhang zwischen den *LV* in Form eines regressionsanalytischen Modells, während das Messgleichungssystem (*MGS*) die Beziehungen zwischen den *LV* und ihren Indikatoren in Form eines faktoren- oder regressionsanalytischen Ansatzes beschreibt.

Ziel ist es, anhand empirischer Daten die Parameter der beiden Gleichungssysteme zu bestimmen und im fachwissenschaftlichen Kontext zu interpretieren.

Zur Lösung dieser Aufgabe haben sich im Wesentlichen zwei Ansätze herauskristallisiert: Die Kovarianzstrukturanalyse und die PLS-Pfadanalyse.

Die Kovarianzstrukturanalyse verfolgt die Idee, die Parameter der Gleichungsmodelle so zu wählen, dass die aus dem Modell theoretisch ableitbare Kovarianzmatrix zwischen den *MV* eine möglichst gute Annäherung an die empirische Kovarianzmatrix erfährt. Dieser Ansatz verlangt keine explizite Kenntnis der *LV*, ist aber insbesondere bei der Modellbewertung auf die Kenntnis von Verteilungseigenschaften der beteiligten Variablen angewiesen.

Einen etwas anderen Weg verfolgt der PLS-Ansatz. Mittels eines zweistufigen Vorgehens werden auf der ersten Stufe konkrete Schätzwerte für die *LV* generiert und auf der zweiten Stufe mit diesen Schätzwerten die Parameter in den beiden Gleichungssystemen geschätzt. Der Vorteil des PLS-Ansatzes besteht vor Allem darin, dass er nur wenige Voraussetzungen an die Verteilungen der beteiligten Variablen stellt. Das liefert allerdings auch den Nachteil fehlender (globaler) Testmöglichkeiten für die Modellbewertung.

Die Kovarianzstrukturanalyse wurde maßgeblich durch die Arbeiten von JÖRESKOG & SÖRBOM (s. z. B. JÖRESKOG (1981) bzw. JÖRESKOG & SÖRBOM (1988)) sowie BENTLER (s. z. B. BENTLER & WEEKS, 1980) vorangebracht. Der PLS-Ansatz wurde von WOLD (s. z. B. 1973 und 1982) bewusst als Alternative zur Kovarianzstrukturanalyse entwickelt.

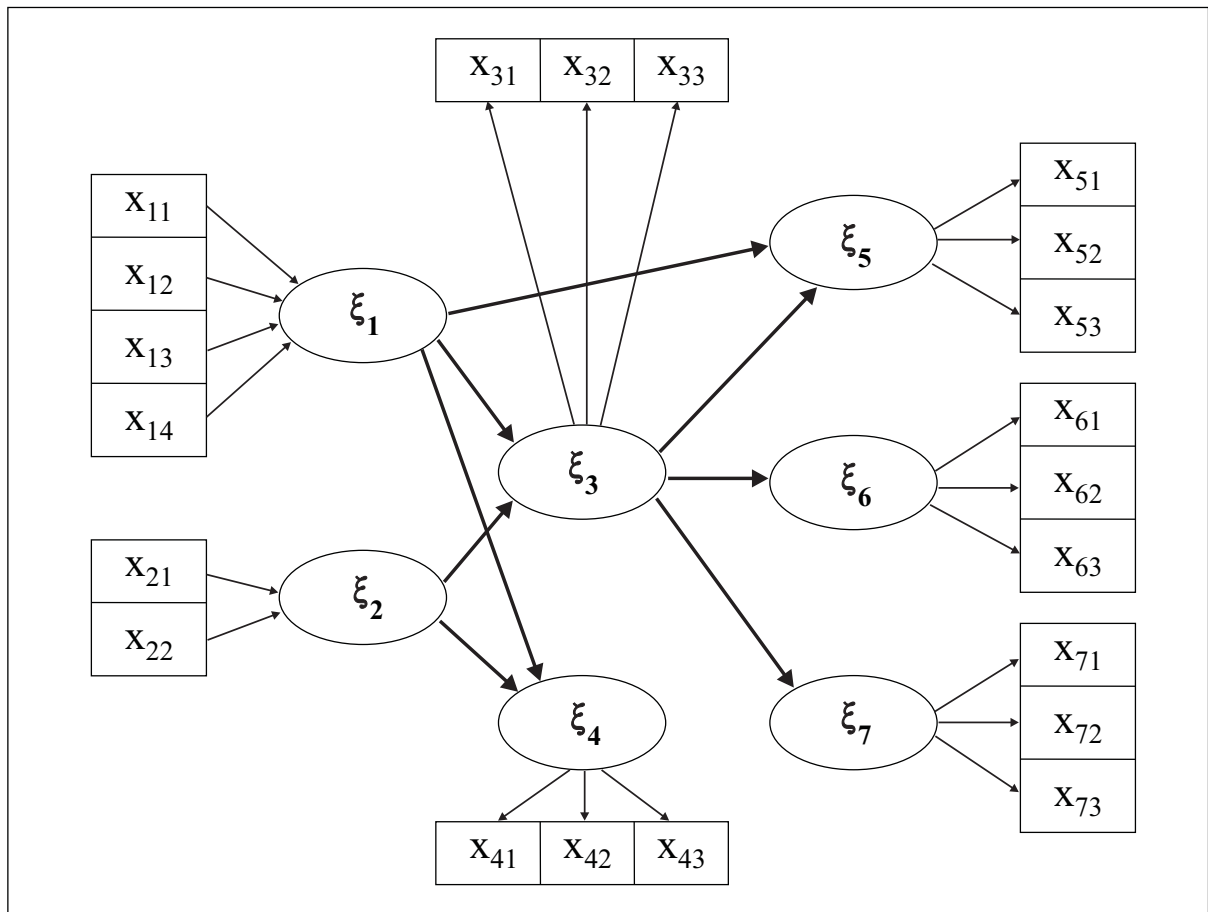


Abbildung 1.1: Beispiel für ein Pfadmodell

Im Folgenden wird die Funktionsweise des Grundmodells (Basismodells) des PLS-Ansatzes erläutert. Als Grundmodell verstehen wir dabei u. a., dass sowohl das Struktur- als auch das Messmodell als Systeme *linearer* Gleichungen angenommen werden.

1.1 Notation

Manifeste Variable werden mit dem Buchstaben x und entsprechenden Indizes belegt. Es wird im folgenden keine Unterscheidung in der Schreibweise für theoretische Variable oder deren empirische Realisationen vorgenommen, aus dem Kontext heraus sollte dies jeweils erkennbar sein. Latente Variable werden mit dem griechischen Buchstaben ξ und ebenfalls entsprechenden Indizes bezeichnet.

Den Ausgangspunkt bilden M Blöcke manifester Variablen, wobei jeder Block $J_{(m)}$ ($m = 1, \dots, M$) Variablen enthalten soll. In runde Klammern gesetzte Indizes beschreiben immer den zugehörigen Variablenblock.

Angelehnt an MATHES (1993a) nehmen wir die folgende Indexnotation vor: Es bezeichne C_m die Menge der Indizes der mit der LV $\xi_{(m)}$ direkt verbundenen LV. Mit C_m^P seien die Indizes der Vorgänger (predecessors) und mit C_m^S die Indexmenge der Nachfolger (successors) bezeichnet. Aufgrund der später geforderten Rekursivität des Modells ergibt sich, dass die Vereinigung beider Indexmengen gerade wieder C_m ergibt und ihr Durchschnitt leer ist. Beispielhaft für das Modell in Abb. 1.1 ergibt sich

$$C_3 = C_3^P \cup C_3^S = \{1, 2\} \cup \{5, 6, 7\}.$$

Weitere Notation werden im Text erklärt.

1.2 Die Modellgleichungen im Pfadmodell

Die wesentlichen Bestandteile von Pfadmodellen sind

- die Zusammengehörigkeit mehrerer manifester Variablen zu so genannten MV-Blöcken,
- die Repräsentation dieser MV-Blöcke durch zugehörige LVE,
- die Darstellung der Beziehungen zwischen MV und ihren LV als lineare Gleichungen und
- eine angenommene kausale Struktur zwischen den LV, die ebenfalls als lineare Gleichungen modelliert werden.

Die Beziehungen im Pfadmodell werden durch zwei Systeme linearer Gleichungen beschrieben. Das *Messmodell* beschreibt die Beziehungen zwischen manifesten und zugehörigen latenten Variablen. Das *Strukturmodell* beschreibt die Beziehungen der latenten Variablen untereinander.

Im Basismodell werden die MV als intervallskalierte, zufällige Variablen vorausgesetzt. Es wird weiter angenommen, dass die MV-Blöcke und ihre Variablen durch jeweils eine latente Variable repräsentiert werden¹.

Wir beschränken uns auf den Fall, dass die MV in standardisierter Form vorliegen, d. h. mit Erwartungswert 0 und Varianz 1. Dies ist keine zwingende Notwendigkeit, erleichtert

¹Es gibt durchaus Modelle, in denen je MV-Block auch mehrere LV zugelassen sind (s. insbesondere MATHES (1993a))

aber die Notation und den weiteren Überblick. Dabei ist die Zentrierung der MV auf den Erwartungswert 0 nicht problematisch, es bedarf lediglich zusätzlicher Lokalisationsparameter in den folgenden Modellgleichungen. Die Annahme gleicher Varianzen aller MV ist aber eine Annahme, die die Interpretation der Parameter grundsätzlich verändert. Daher ist durch den Fachwissenschaftler immer zu prüfen, ob nicht unterschiedliche Varianzen (also Skalierungen der MV) im Pfadmodell erhalten bleiben müssen.

Die im Pfadmodell angenommenen kausalen Beziehungen zwischen den LV werden in einem Strukturgleichungssystem (*SGS*) als multiples lineares Regressionsmodell dargestellt. Die jeweilige latente Variable $\xi_{(m)}$ erscheint dabei als Regressand ihrer im Pfadmodell auftretenden direkten Vorgänger $\xi_{(l)}$:

SGS:

$$\xi_{(m)} = \sum_{l \in C_{(m)}^P} \xi_{(l)} \gamma_{(ml)} + \varepsilon_{(m)}, \quad m = 1, \dots, M, \quad (1.1)$$

bzw. mit $\Xi = (\xi_{(1)}, \dots, \xi_{(M)})$ in Matrixschreibweise

$$\xi = \Xi \cdot \Gamma + \varepsilon. \quad (1.2)$$

Die Koeffizienten $\gamma_{(ml)}$ werden als Pfadkoeffizienten bezeichnet. Die $\varepsilon_{(m)}$ stellen Fehlergrößen dar, von denen erwartet wird, dass sie den Erwartungswert 0 besitzen².

Im Basismodell wird weiter angenommen, dass das Strukturmodell ein *rekursives* Modell ist, d. h. es gibt keine (direkten oder indirekten) Beziehungen einer LV auf sich selbst. Das führt dazu, dass die Koeffizientenmatrix Γ – eventuell mittels Umordnung der LV – als untere bzw. obere Dreiecksmatrix mit Nullen auf der Hauptdiagonale dargestellt werden kann.

Die im Pfadmodell angenommenen kausalen Beziehungen zwischen MV und zugehörigen LV werden in einem Messgleichungssystem (*MGS*) dargestellt. Innerhalb der einzelnen Variablenblöcke entscheidet dann die unterstellte Wirkrichtung über die konkrete Form des Messmodells. Es wird zwischen reflektiven und formativen Messmodellen unterschieden (siehe auch den vorangegangenen Beitrag):

Reflektives MGS: Sind die MV als Wirkung der zugehörigen LV modelliert, also als Reflektionen der LV, wird das Messmodell in dem entsprechenden Variablenblock als

²Eine detaillierte Modellbeschreibung insbesondere der Regressionsgleichungen soll hier aus Platzgründen nicht ausgeführt werden. Der geeignete Leser sei dazu an FORNELL & CHA (1994) verwiesen.

faktorenanalytisches Modell dargestellt:

$$x_{(m)j} = \xi_{(m)}\lambda_{(m)j} + \theta_{(m)j}, \quad m = 1, \dots, M; j = 1, \dots, J_{(m)}, \quad (1.3)$$

bzw. in Matrixschreibweise:

$$X_{(m)} = \xi_{(m)}\lambda_{(m)} + \theta_{(m)}. \quad (1.4)$$

Formatives MGS: Sind die MV als ursächlich für die Bildung der LV angenommen – formen sie also die LV – wird das Messmodell in dem entsprechenden Variablenblock als multiples (lineares) Regressionsmodell dargestellt:

$$\xi_{(m)} = \sum_{j=1}^{J_{(m)}} x_{(m)j}\omega_{(m)j} + \theta_{(m)} \quad m = 1, \dots, M, \quad (1.5)$$

bzw. in Matrixschreibweise:

$$\xi_{(m)} = X_{(m)}\omega_{(m)} + \theta_{(m)}. \quad (1.6)$$

Die manifesten Variablen werden auch als Indikatoren der latenten Variablen und die Koeffizienten $\lambda_{(m)j}$ im reflektiven Messmodell als Ladungen bzw. die Koeffizienten $\omega_{(m)j}$ im formativen Messmodell als Gewichte bezeichnet. In $\theta_{(m)}$ sind die jeweiligen Messfehlervariable zusammengefasst.

2 Der PLS-Ansatz zur Schätzung der latenten Variablen

Die Schätzung der LV geschieht mit Hilfe eines zusätzlich einzuführenden Gewichtungsgleichungssystems (GGs), welches die LV als gewichtete Summe der MV *definiert* und damit zumindest formal Verwandtschaft zur Hauptkomponentenanalyse bzw. zur Kanonischen Korrelationsanalyse herstellt.

GGs:

$$\xi_{(m)} = X_{(m)}\omega_{(m)} \quad m = 1, \dots, M. \quad (2.1)$$

Das GGs ist kein stochastisches Modell: „*Ein wesentlicher Unterschied zwischen PLS und LISREL betrifft die Zusatzannahmen, die das Modell schätzbar machen. In PLS werden*

die latenten Variablen als gewichtete Aggregate geschätzt; das impliziert, dass die LV-Werte (Faktorwerte) zum konstitutiven Modellbestandteil werden“ (LOHMÖLLER, 1984a, S. 45).

Die Gewichtungskoeffizienten

$$\Omega_{(m)} = (\omega_{(m)1}, \omega_{(m)2}, \dots, \omega_{(m)J_{(m)}})'$$

zur Bestimmung der LV in dem jeweiligen MV-Block sind zunächst unbekannt und werden in einem Iterationszyklus (s. u.) unter bestimmten Optimierungskriterien geschätzt. Auf die einzelnen MV-Blöcke bezogen ist das PLS-Modell durch das GGS mit der Hauptkomponentenanalyse verwandt. Während allerdings in der Hauptkomponentenanalyse die Gewichte ausschließlich Informationen der MV tragen, muss hier die innere Struktur des Pfadmodells, die Beziehungen der LV untereinander, Eingang in die Bestimmung der Gewichtungskoeffizienten finden. Zu diesem Zwecke werden sogenannte Umgebungsvariable (UV, engl.: *instrument variables*) definiert:

UV:

$$\xi_{(m)}^* = \sum_{l \in C_m} \xi_{(l)} \rho_{(ml)} \quad m = 1, \dots, M. \quad (2.2)$$

Die in (2.2) auftretenden Koeffizienten $\rho_{(ml)}$ sind Gewichtungskoeffizienten, auf die im Abschnitt 2.3 näher eingegangen wird und die zu unterschiedlichen Optimalitätseigenschaften des PLS-Modells führen. Die Umgebungsvariablen dienen als Approximation der gesuchten LVen aus dem Strukturmodell heraus (s. 1.1). Zu beachten ist, dass in die UV $\xi_{(m)}^*$ die Beziehungen zu allen mit $\xi_{(m)}$ benachbarten LV, also sowohl mit den Vorgängern als auch mit den Nachfolgern, eingehen. Unter Einbeziehung des GGS und der UV ist das PLS-Modell „somit eine Generalisierung der Hauptkomponentenanalyse ...“ (LOHMÖLLER, 1984a, S. 45). Vielleicht sollte man es eher als Verallgemeinerung der kanonischen Korrelationsanalyse bezeichnen, da zwar die LV mit $\xi_{(m)} = X_{(m)}\omega_{(m)}$ eine Linearkombination der MV darstellen, allerdings die Gewichte (s. u.) so gewählt werden, dass sie, in einem noch zu spezifizierenden Sinne, möglichst gut zu den Umgebungsvariablen $\xi_{(m)}^*$ passen. Die Abbildung 2.1 liefert eine ungefähre Vorstellung der Einflüsse auf die jeweilige LV.

Zur Schätzung der LV-Werte wird im PLS-Algorithmus eine Kleinste-Quadrat-Methodik zur Bestimmung von Hauptkomponenten auf das Problem mehrerer Variablenblöcke erweitert und zusätzlich berücksichtigt, dass in einem Pfadmodell nicht jeder MV-Block mit jedem anderen MV-Block verbunden ist. Zum besseren Verständnis der Methodik werden wir ausgehend vom Problem der PCA (Prinzipal Component Analysis; Hauptkom-

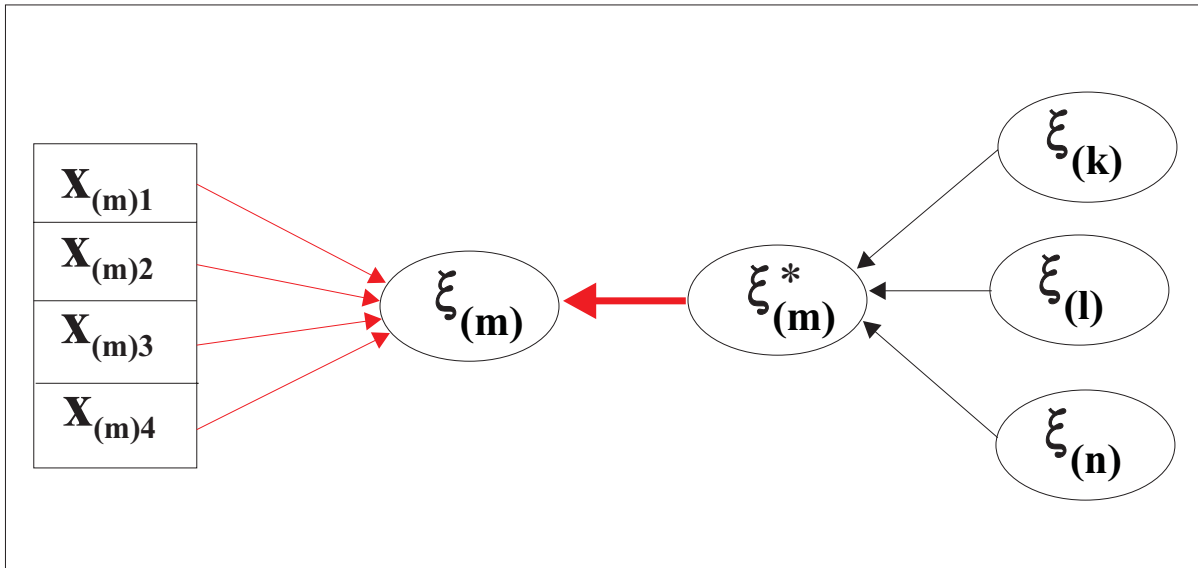


Abbildung 2.1: Beeinflussung der LV $\xi_{(m)}$ sowohl durch die zugehörigen MV als auch durch die benachbarten LV.

ponentenanalyse) den Schätzalgorithmus auf Probleme der verallgemeinerten kanonischen Korrelation erweitern und schließlich die Anwendung im Pfadmodell zeigen.

2.1 Der ALS Algorithmus zur Bestimmung der ersten Hauptkomponente

Das Problem der Bestimmung der ersten Hauptkomponente in der PCA bedeutet die Suche nach einer (latenten) Variablen ξ als Linearkombination von beobachteten Variablen $X = (x_1, \dots, x_J)$ (einem MV-Block)

$$\xi = X \cdot \omega,$$

die als „Ersatz“ für die J (manifesten) Variablen x_j dienen und dabei möglichst viel Information der ursprünglichen Variablen behalten soll. Möglichst viel Information wird dabei in einem geometrischen Sinne interpretiert, indem gefragt wird, wie stark die quadrierte Abweichung der neuen Variable ξ von den gewichteten Ursprungsvariablen insgesamt ist. Dies wird in einer sogenannten Kleinste-Quadrat-Verlustfunktion dargestellt:

$$\sigma(\xi; X, \omega) := \sum_{j=1}^J \|\xi - x_j \omega_j\|^2 \quad (2.3)$$

und es werden solche Gewichte ω gesucht, die diese „Verlustfunktion“ minimieren¹. Für realisierte Stichprobenwerte in X liefert der Mittelwert

$$\xi = \frac{1}{J} \sum_{j=1}^J X_j \omega_j$$

das gewünschte Minimum für gegebene Werte von ω , d. h. die Definition von ξ als Linearkombination der Variablen in X ist eine Folge der gewählten quadratischen Verlustfunktion.

Bezüglich ω wird die Verlustfunktion (2.3) minimal, indem ω identisch gleich 0 gesetzt wird. Dann ist auch ξ identisch Null und der „Informationsverlust“ scheint ebenfalls Null zu sein, was natürlich nicht stimmt. Um diese triviale Lösung zu verhindern bedarf es einer Normierung, die hier so vorgenommen werden soll, dass $\|\xi\|^2 = 1$ gesetzt wird.

Unter diesen Umständen, ausgehend von einer beliebigen Anfangsbelegung für $\omega_{\langle 0 \rangle} \neq \mathbf{0}$ lässt sich folgender Algorithmus installieren (vgl. z. B. GIF1, 1990, S. 84 ff.)²:

- (ALS1) $\xi_{\langle k+1 \rangle} := \frac{1}{J} X \cdot \omega_{\langle k \rangle} \cdot c$ ξ als gewogenes arithmetisches Mittel der MV
 $c := \frac{1}{J} \left(\omega'_{\langle k \rangle} X' X \omega_{\langle k \rangle} \right)^{-1/2}$ Normierungsfaktor zur Normierung der ξ
- (ALS2) $\omega_{\langle k+1 \rangle} := X' \xi_{\langle k+1 \rangle}$ Berechnung neuer Gewichte ω
Konvergenztest

Dieser Alternating Least Squares Algorithmus (ALS; alternierende kleinste Quadrate) besteht darin, die Aufgabe

$$\xi = X \cdot \omega$$

für die obige Verlustfunktion (2.3) für zunächst unbekannte ξ und ω in zwei Teilaufgaben zu zerlegen und diese alternierend zu lösen.

Das Verfahren ist aus der Literatur als *Potenzmethode* (engl. „power method“) oder *von Mises Methode* bekannt und lässt sich als Zyklus wie in Abb. 2.2 darstellen. Die Konvergenz des Algorithmus ist gesichert, solange die Matrix $X'X$ einen dominanten Eigenwert besitzt³, was in praktischen Fällen mit Wahrscheinlichkeit 1 immer der Fall ist, solange mehr Fälle als Variable auftauchen und nicht zwei genau linear abhängige Variable in der Matrix X erscheinen.

Bezogen auf ein Pfadmodell liefert die PCA diejenigen LV, die am besten im Sinne der obigen Verlustfunktion zu ihren manifesten Indikatoren passt.

¹Dies liefert gleichzeitig eine LV, die als Linearkombination der MV einen maximalen Anteil an der Gesamtvarianz von X (Summe der Varianzen der einzelnen x_j) ausweist.

²Der Index $\langle k \rangle$ ist hier wie im folgenden ein Iterationszähler.

³Auf mathematische Details soll hier nicht näher eingegangen werden.

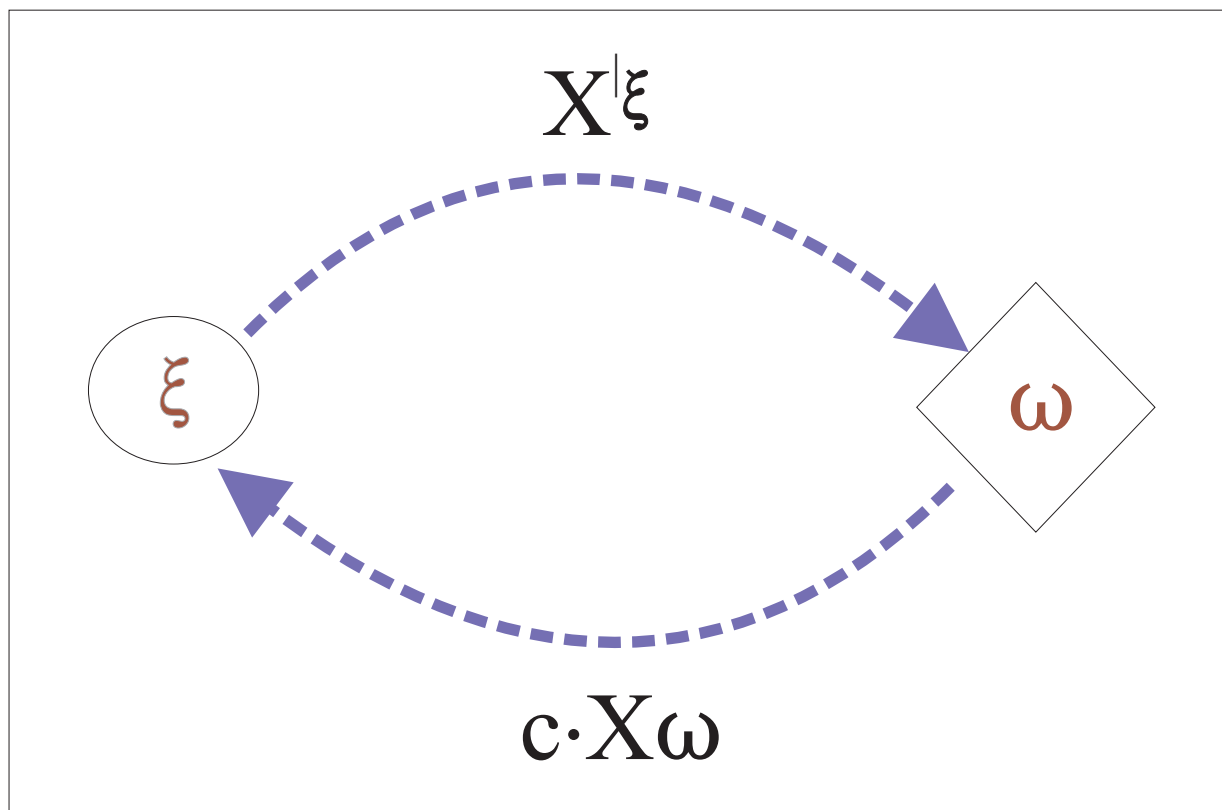


Abbildung 2.2: ALS-Algorithmus in der PCA.

Erweitert man die Fragestellung auf mehrere Blöcke manifester Variabler, die jeweils durch eine LV repräsentiert werden sollen, wobei diese LV zusätzlich möglichst hoch miteinander korrelieren sollen, ergibt sich das Problem der Kanonischen Korrelationsanalyse.

2.2 Der ALS-Algorithmus in der Kanonischen Korrelationsanalyse

Die Aufgabenstellung in der Kanonischen Korrelationsanalyse (CCA, Canonical Correlation Analysis) ist insofern eine Erweiterung des Konzeptes der PCA, als dass jetzt mehrere MV-Blöcke betrachtet werden, für die einzelne LV angegeben werden sollen. Neben der Forderung, dass die LV Linearkombinationen der zugehörigen MV sein sollen, werden zusätzliche Anforderungen gestellt, die den Zusammenhang der LV untereinander betreffen. Das klassische Modell der kanonischen Korrelationsanalyse geht von zwei MV-Blöcken $X_{(1)}$ und $X_{(2)}$ aus und verlangt, dass die zugehörigen LV $\xi_{(1)} = X_{(1)}\omega_{(1)}$ und $\xi_{(2)} = X_{(2)}\omega_{(2)}$ so gewählt werden, dass die Korrelation zwischen den beiden LV maximal wird. Dies lässt

sich in eine allgemeinere Aufgabe der Minimierung folgender Verlustfunktion überführen

$$\sigma(\omega_{(1)}, \dots, \omega_{(M)}) = \sum_{m=1}^M \|X_{(m)}\omega_{(m)} - \xi\|^2, \quad (2.4)$$

wobei für gegebene $\omega_{(m)}$ das Konstrukt ξ wiederum das arithmetische Mittel der $X_{(m)}\omega_{(m)}$ ist. Bezeichnen wir mit $\xi_{(m)} := X_{(m)}\omega_{(m)}$ die entsprechenden Linearkombinationen in den MV-Blöcken, so erhalten wir mit $\xi = \frac{1}{M}\xi_{(m)}$

$$\begin{aligned} \sum_{m=1}^M \|X_{(m)}\omega_{(m)} - \xi\|^2 &= \sum_{m=1}^M \omega'_{(m)} X'_{(m)} X_{(m)} \omega_{(m)} + \left(M\xi'\xi - 2\xi' \sum_{m=1}^M X_{(m)}\omega_{(m)} \right) \\ &= \sum_{m=1}^M \omega'_{(m)} X'_{(m)} X_{(m)} \omega_{(m)} - \frac{1}{M} \sum_{m=1}^M \sum_{l=1}^M \omega'_{(m)} X'_{(m)} X_{(l)} \omega_{(l)} \\ &= \sum_{m=1}^M \|\xi_{(m)}\|^2 - \frac{1}{M} \sum_{m=1}^M \sum_{l=1}^M \xi'_{(m)} \xi_{(l)}. \end{aligned}$$

Sind die latenten Variablen normiert, so erscheint wegen $\xi'_{(m)}\xi_{(l)} = \text{cor}(\xi_{(m)}, \xi_{(l)})^4$ gerade die Summe der Korrelationskoeffizienten mit negativem Vorzeichen in der Verlustfunktion. Die Verlustfunktion minimieren heißt dann gerade die Summe der Korrelationskoeffizienten zu maximieren. Für den Fall $M = 2$ heißt das aber gerade genau die Linearkombinationen zu finden, die maximal miteinander korrelieren.

Ähnlich wie in der PCA versuchen wir, eine „gemeinsame“ LV zu finden, die die einzelnen LV in den Blöcken mit möglichst wenig Informationsverlust beschreibt. Hier ist allerdings die Beziehung der einzelnen LV zu ihren MV nicht mehr dadurch geprägt, dass sie einen möglichst hohen Anteil erklärter Varianz in dem zugehörigen MV-Block tragen muss, wie das in der PCA der Fall ist. Aber wie in der PCA lässt sich auch hier ein Iterationsalgorithmus mit alternierenden Kleinste-Quadrat-Methoden angeben, allerdings nicht mehr mit Normierung der gemeinsamen LV ξ , sondern Normierung der LV $\xi_{(m)}$ in den MV-Blöcken⁵. Beschreiben wir mit ω den Gesamtgewichtsvektor und mit

$$D_{X'X} := \begin{pmatrix} X'_1 X_1 & 0 & \vdots & 0 \\ 0 & X'_2 X_2 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & X'_M X_M \end{pmatrix}$$

⁴Mit $\text{cor}(x, y)$ bezeichnen wir den Korrelationskoeffizienten der beiden beteiligten Variablen.

⁵Normierung der gemeinsamen LV ξ ist dann ein Spezialfall und also auch enthalten.

die Blockdiagonalmatrix bzgl. der MV-Blöcke, so lautet der Algorithmus:

$$\begin{aligned}
 (\text{ALS1}) \quad \xi &:= \frac{1}{M} X \omega c = \frac{1}{M} \sum_{m=1}^M X_{(m)} \omega_{(m)} c && \xi \text{ als gewogenes arithm. Mittel der MV} \\
 c &:= \frac{1}{M} (\omega' D_{X'X} \omega)^{-1/2} && \text{Faktor zur Normierung der } \xi_{(m)} \\
 (\text{ALS2}) \quad \omega &:= X' \xi && \text{Berechnung neuer Gewichte } \omega \\
 &&& \textit{Konvergenztest}
 \end{aligned}$$

Im Schätzalgorithmus des PLS-Modells zur Schätzung der LV wird im Folgenden versucht, sowohl die Aspekte der PCA (möglichst guter Zusammenhang zwischen LV und MV) als auch die Aspekte der CCA (möglichst hohe Korrelationen zwischen den LV) zu berücksichtigen und zusätzlich zu beachten, dass ein bestimmtes Pfaddesign dazu führt, dass nicht mehr jeder MV-Block mit jedem anderen MV-Block über eine LV verbunden ist, wie das in der CCA der Fall ist.

2.3 Der PLS Schätzalgorithmus

Der von H. WOLD (s. z. B. 1966b, 1973) Ende der 60er, Anfang der 70er Jahre eingeführte Basisalgorithmus zur Berechnung der latenten Variablen innerhalb eines Pfadmodells berechnet für jeden einzelnen Variablenblock $X_{(m)}$ in $X = (X_{(1)}, \dots, X_{(M)})$ schrittweise neue latente Variable $\xi_{(m)}$ unter der Annahme, dass die LV der anderen Variablenblöcke $X_{(l)}$ ($l \neq m$) bekannt sind. Dies liefert den Namensbestandteil *partial* am Namen *Partial Least Squares*. Formal ist der Algorithmus wie folgt darstellbar:

Nach der Bereitstellung von (beliebigen) Startgewichten $\omega_{(m)(0)} \neq \mathbf{0}$ ($m = 1, \dots, M$) werden erste Approximationen der LV bereitgestellt

$$\xi_{(m)(0)} = X_{(m)} \omega_{(m)(0)} \cdot c_{(m)(0)} \quad (m = 1, \dots, M)$$

und anschließend der folgende Iterationsalgorithmus durchgeführt:

(PLS1): Bestimmung von Umgebungsvariablen⁶

$$\xi_{(m)(k)}^* = \frac{1}{|C_m|} \sum_{l \in C_m} \xi_{(l)(k)} \rho_{(ml)(k)} \quad (2.5)$$

als gewichtete Summe benachbarter LV.

⁶Die Umgebungsvariable ist im Originalalgorithmus von Wold nicht mit dem Faktor $\frac{1}{|C_m|}$ versehen. Die im Weiteren erfolgende Normierung der LV liefert dann aber wieder dieselben Ergebnisse.

(PLS2): Bestimmung von Gewichtsvektoren

$$\begin{aligned} \text{Mode A:} \quad & X_{(m)} = \xi_{(m)\langle k \rangle}^* \omega'_{(m)\langle k+1 \rangle} + U_{(m)} \\ \text{Mode B:} \quad & \xi_{(m)\langle k \rangle}^* = X_{(m)} \omega_{(m)\langle k+1 \rangle} + u_{(m)} \end{aligned} \quad (2.6)$$

zur Berechnung neuer Approximationen der LV im Messmodell.

(PLS3): Bestimmung neuer Werte für die LV als Linearkombinationen aus den zugehörigen MV

$$\xi_{(m)\langle k+1 \rangle} = X_{(m)} \omega_{(m)\langle k+1 \rangle} c_{(m)\langle k+1 \rangle} \quad (2.7)$$

mit dem Normierungsfaktor $c_{(m)\langle k+1 \rangle} = \left(\omega'_{(m)\langle k+1 \rangle} X'_{(m)} X_{(m)} \omega_{(m)\langle k+1 \rangle} \right)^{-\frac{1}{2}}$.

Konvergenztest:

$$\sum_{m=1}^M \|\xi_{(m)\langle k+1 \rangle} - \xi_{(m)\langle k \rangle}\|^2 < \varepsilon$$

Im Schritt (PLS1) sind bereits Näherungen der $\xi_{(m)}$ bekannt, die schrittweise an das Pfadmodell angepasst werden sollen. Der Iterationsalgorithmus wird in jedem Zyklus (PLS1-PLS3) für jeden Block separat durchlaufen und erst anschließend der nächste Iterationszyklus in Angriff genommen⁷. Nach jedem Iterationszyklus stehen also für jeden MV-Block neue Gewichte $\omega_{(m)\langle k \rangle}$ und damit neue latente Variablen $\xi_{(m)\langle k \rangle}$ zur Verfügung.

Die Anpassung der $\xi_{(m)}$ erfolgt wieder in einem alternierenden Algorithmus. Zunächst werden (in PLS1) sogenannte Umgebungsvariable (s. (2.2)) als gewichtete Summe benachbarter LV definiert. Dies entspricht im wesentlichen dem Schritt (ALS1) der CCA und versucht LV zu definieren, die möglichst hohe Korrelationen untereinander aufweisen. Im Unterschied zur CCA, wo für jeden MV-Block im Prinzip die gleiche Umgebungsvariable (mit allen anderen LV)

$$\xi_{(m)}^* = \sum_{l \neq m} X_{(l)} \omega_{(l)} \cdot c$$

zur Verfügung steht⁸, da jeder Block mit jedem anderen verbunden ist, brauchen hier die einzelnen Blöcke jeweils unterschiedliche Umgebungsvariablen. Die Umgebungsvariablen dienen als Näherung der LV aus dem Strukturmodell heraus. Im folgenden Schritt

⁷Ein von WOLD (1985) vorgeschlagener Algorithmus aktualisiert die jeweils berechneten LV sofort für den nächsten Iterationsschritt. Es ist bisher noch unklar, ob das zu gleichen Ergebnissen führt, es spricht allerdings einiges dafür, dass dies zumindest häufig der Fall ist.

⁸Benutzt man genau diese Umgebungsvariable im PLS-Algorithmus und keine Gewichtung ($c = 1$) so führt die iterative Bestimmung der $\eta_{(m)}$ nach Modus B zu genau den Ergebnissen des ALS Algorithmus für die CCA.

(PLS2) werden Gewichtsvektoren im Sinne der PCA gesucht (PCA-ALS2), die gewährleisten sollen, dass die Umgebungsvariablen möglichst gut durch die MV im zugehörigen Variablenblock repräsentiert werden. Dabei muss allerdings berücksichtigt werden, ob die Beziehung $MV \leftrightarrow LV$ reflektiver oder formativer Natur ist. Dies geschieht durch Auswahl des für den jeweiligen MV-Block zutreffenden Modus A oder B.

Die Gewichtsvektoren werden im abschließenden Schritt (PLS3) benutzt, um die LV als Linearkombinationen der ihnen zugehörigen MV zu definieren.

Damit sind neue LV in den Variablenblöcken gefunden, die zum Einen die innere Struktur über die Umgebungsvariablen (PLS1) und zum Anderen die äußere Struktur des Messmodells über die Bestimmung der Gewichte aus dem Zusammenhang zwischen Umgebungsvariable und MV (PLS2) widerspiegeln.

Eine etwas bildhaftere Vorstellung des Iterationszyklus lässt sich der Abb. 2.3 entnehmen.

Gewichtungskoeffizienten zur Bestimmung der Umgebungsvariablen

Die Umgebungsvariable in einem MV-Block ist definiert als

$$\xi_{(m)}^* = \frac{1}{|C_m|} \sum_{l \in C_m} X_{(m)} \omega_{(m)} \cdot \rho_{(ml)} = \frac{1}{|C_m|} \sum_{l \in C_m} \xi_{(l)} \cdot \rho_{(ml)}. \quad (2.8)$$

Entgegen dem SGS, indem die LV selbst jeweils nur als Regressand ihrer Vorgänger auftritt ($l \in C_m^P$) ist die UV eine Linearkombination aller direkt benachbarten LV ($l \in C_m$). Dieses Vorgehen ist notwendig, um die Gewichte zur Errechnung der LV-Scores (im Schritt PLS2) so zu bestimmen, dass sie auch gut in das innere Modell passen. Die Idee ist, die Koeffizienten $\rho_{(ml)}$ so zu wählen, dass sie die bivariaten Zusammenhänge der LV zu ihren Nachbarn widerspiegeln. Im Basisalgorithmus werden entweder die Korrelationskoeffizienten zwischen den LV verwendet (PLS1b, Korrelationsgewichtung, s. u.) oder deren Vorzeichen (PLS1a, Vorzeichengewichtung), was im Zusammenhang mit den unterschiedlichen Modi (s. PLS2) zu unterschiedlichen Optimierungseigenschaften innerhalb des Pfadmodells führt.

Vorzeichengewichtung (PLS1a): Bei der Vorzeichengewichtung bezeichnet der innere Gewichtungsfaktor

$$\rho_{(ml)} = \begin{cases} \operatorname{sgn}(\operatorname{cor}(\xi_{(m)}, \xi_{(l)})) & \text{für } l \in C_m \\ 0 & \text{sonst} \end{cases} \quad (2.9)$$

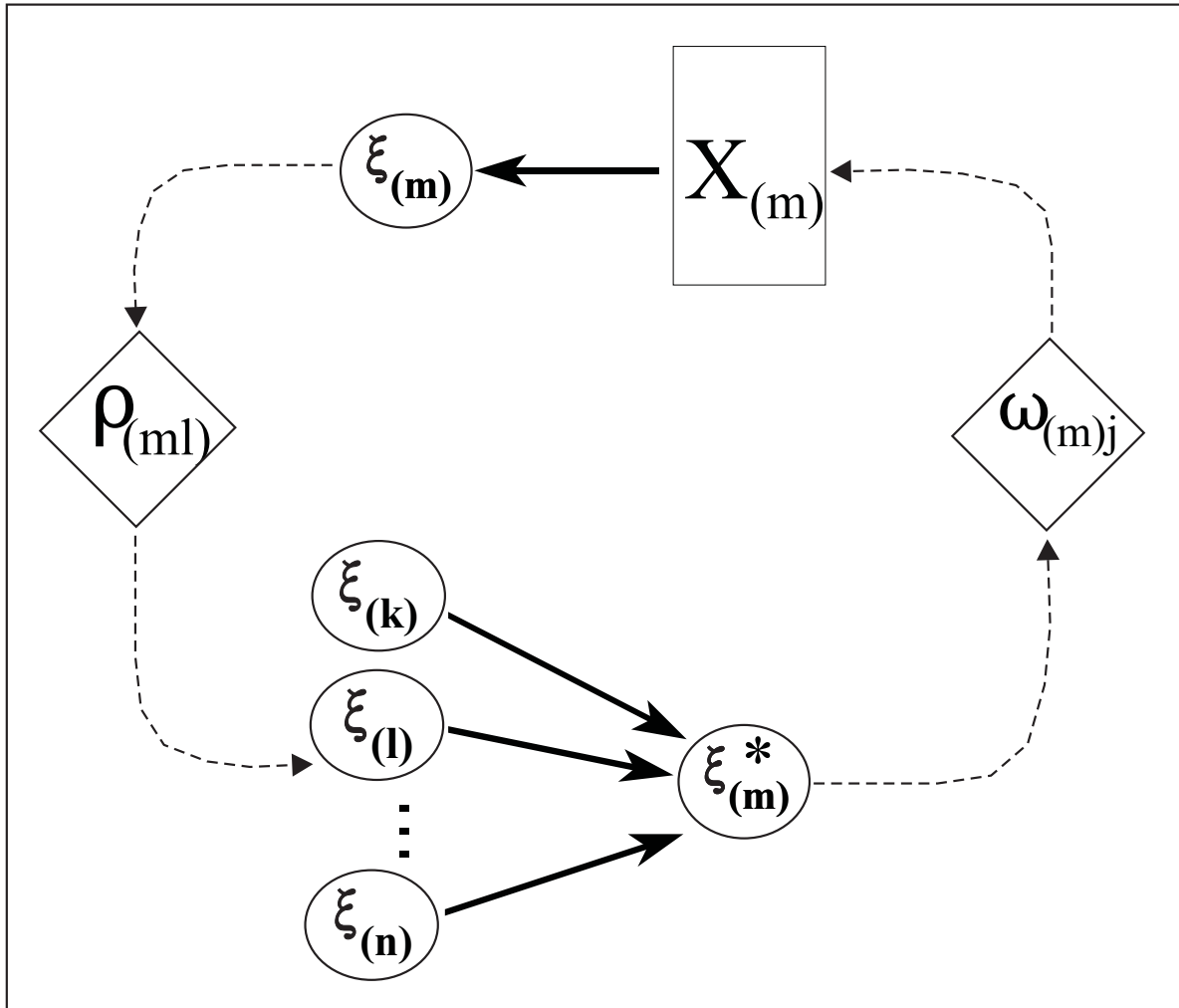


Abbildung 2.3: Graphische Interpretation des PLS Iterationsalgorithmus.

das Vorzeichen des Korrelationskoeffizienten von im Pfadmodell direkt miteinander verbundenen LV, bei nicht miteinander verbundenen LV ist er Null. Damit wird erstens gesichert, dass nur zusammenhängende Blöcke in die spätere Analyse eingehen ($\rho_{(ml)} = 0$ sonst), und zweitens wird zumindest die Richtung (das Vorzeichen) des Einflusses mitberücksichtigt.

Korrelationsgewichtung (PLS1b): Bei der Korrelationsgewichtung berücksichtigt der innere Gewichtungsfaktor

$$\rho_{(ml)} = \begin{cases} \text{cor}(\xi_{(m)}, \xi_{(l)}) & \text{für } l \in C_{(m)} \\ 0 & \text{sonst} \end{cases} \quad (2.10)$$

neben der Richtung auch die Stärke der Korrelation der jeweils eingehenden LV.

Mit der Wahl der Korrelationskoeffizienten als Gewichte der LV wird für die einzelnen LV das folgende Regressionsproblem gelöst:

$$\xi_{(m)} = \xi_{(l)}c_{ml} + \varepsilon_m,$$

d. h. wie müssen die benachbarten LVen skaliert werden, damit sie $\xi_{(m)}$ möglichst gut vorhersagen können? Anders ausgedrückt finden wir die Korrelationsgewichte als Lösung der Optimierungsaufgabe

$$\min_c \sum_{l \in C_m} \|\xi_{(m)} - c_{ml}\xi_{(l)}\|.$$

Für diese reskalierten LV $\rho_{(ml)}\xi_{(l)}$ bestimmen wir dann einen Schwerpunkt, nämlich gerade die UV⁹

$$\xi_{(m)}^* := \arg \min_{\xi} \sum_{l \in C_m} \|\xi - \rho_{(ml)}\xi_{(l)}\|^2,$$

also eine Variable, die im Mittel am besten im Sinne von Korrelationskoeffizienten zu ihrer Umgebung passt. Bei der folgenden Bestimmung der Gewichtsvektoren in (PLS2) für die neue LVe $\xi_{(m)}$ wird gerade eine Anpassung an diesen Schwerpunkt angestrebt. D. h. die neue LV ist aus den MV heraus so zu bestimmen, dass sie zu ihren aktuellen Nachbarn möglichst gut passt. Abhängig von dem gewählten Modus wird dabei die Zusammenhangsstruktur in dem entsprechenden MV-Block $X_{(m)}$ berücksichtigt (Modus A) oder nicht (Modus B).

Wird statt der Korrelationskoeffizienten nur deren Vorzeichen gewählt, so wird zumindest versucht, allen benachbarten LV dieselbe allgemeine Richtung zu geben.

Die Gewichtskoeffizienten zur Berechnung der LV

Für die Berechnung der Gewichte im GGS (2.1) wird angenommen, dass die UV Approximationen der entsprechenden LV aus dem Strukturmodell heraus sind. Die Gewichte werden dann dem jeweiligen Messmodell adäquat bestimmt. Dazu werden die Regressionsgleichungen des Messmodells mit den UVen gelöst und die berechneten Regressionskoeffizienten als Gewichte benutzt.

⁹Man könnte auch versuchen, eine PCA für die benachbarten LV durchzuführen

$$\tilde{\xi}_{(m)} := \arg \min_{\xi} \sum_{l \in C_m} \|\xi - c_{ml}\xi_{(l)}\|^2$$

und die Hauptkomponente $\tilde{\xi}_{(m)}$ als Approximation für die zu bildende neue LV $\xi_{(m)}$ zu benutzen. Allerdings wird mit der Hauptkomponente nicht der Zusammenhang zum MV-Block $X_{(m)}$ berücksichtigt, der aber wesentlicher Bestandteil des Strukturgleichungsmodells ist.

(PLS2) Gewichtungsberechnung:

Modus A: MV-Block mit reflektivem Messmodell (s. Formeln (1.3), (1.4))

$$X_{(m)} = \xi_{(m)\langle k \rangle}^* \omega'_{(m)\langle k+1 \rangle} + U_{(m)} \quad (2.11)$$

bzw.

Modus B: MV-Block mit formativem Messmodell (s. Formeln (1.5), (1.6))

$$\xi_{(m)\langle k \rangle}^* = X_{(m)} \omega_{(m)\langle k+1 \rangle} + u_{(m)}. \quad (2.12)$$

Im Rahmen einer OLS-Regression ergeben sich für die zu bestimmenden Gewichte $\omega_{(m)}$ dann folgende Lösungen:

Modus A:

$$\omega_{(m)\langle k+1 \rangle} = \mathbf{X}'_{(m)} \xi_{(m)\langle k \rangle}^* \left(\xi_{(m)\langle k \rangle}^{*'} \xi_{(m)\langle k \rangle}^* \right)^{-1} \quad (2.13)$$

bzw.

Modus B:

$$\omega_{(m)\langle k+1 \rangle} = \left(X'_{(m)} X_{(m)} \right)^{-1} X'_{(m)} \xi_{(m)\langle k \rangle}^* \quad (2.14)$$

Beide Modi führen zu unterschiedlichen Optimierungseigenschaften im Pfadmodell und zu den unterschiedlichen Annahmen über reflektive (Modus A) und formative (Modus B) Definitionen der MV-Blöcke.

Im *Modus A* werden die Gewichte $\omega_{(m)}$ mit Hilfe von einzelnen univariaten Regressionsgleichungen bestimmt. Bilden wir die neue LV

$$\begin{aligned} \xi_{(m)} &= X_{(m)} \omega_{(m)} \\ &= \|\xi_{(m)}^*\|^2 \sum_{j=1}^{J_m} \frac{X'_{(m)j} \xi_{(m)}^*}{\|\xi_{(m)}^*\|^2 \|X_{(m)j}\|^2} X_{(m)j} \end{aligned}$$

und bezeichnen wir mit $\mathbb{P}_{X_{(m)j}} \left(\xi_{(m)}^* \right)$ die Projektion von $\xi_{(m)}^*$ auf die Achse $X_{(m)j}$ so ist die neue LV $\xi_{(m)}$ gerade der Schwerpunkt dieser Projektionen auf die Achsen von $X_{(m)}$.

Im Unterschied dazu wird im *Modus B* die multiple Regressionsgleichung (2.12) gelöst. Das entspricht der Projektion von $\xi_{(m)}^*$ direkt in den Raum $X_{(m)}$. Hier spielt die Lage der Achsen

$X_{(m)_j}$, die Kovarianzstruktur innerhalb des MV-Blockes, keine Rolle mehr. Ausgedrückt in Minimierungsproblemen erhalten wir für *Modus A*

$$\omega_{(m)} := \arg \min_{\omega} \sum_{j=1}^{J_m} \left\| \xi_{(m)}^* - \omega_j X_{(m)_j} \right\|^2$$

bzw. für *Modus B*

$$\omega_{(m)} := \arg \min_{\omega} \left\| \xi_{(m)}^* - \sum_{j=1}^{J_m} \omega_j X_{(m)_j} \right\|^2.$$

Während also im *Modus B* keine Rücksicht auf die Lage der einzelnen Achsen genommen werden muss – es ist ein Gesamtkonstrukt $\sum_{j=1}^{J_m} \omega_j X_{(m)_j}$ anzupassen – wird im *Modus A* jede einzelne Achse an die UV angepasst und damit die Struktur innerhalb des MV Blockes von Bedeutung für die Bildung der neuen LV $\xi_{(m)}$. Das liefert zumindest heuristisch eine Begründung für die Optimierungseigenschaften der beiden Modi.

Modus A: Vorzeichengewichtung: Der PLS Algorithmus liefert ein Maximum für das Produkt aus erklärter Varianz im MV Modell und Beträgen der Korrelationskoeffizienten der LV im Strukturmodell. Die LV $\xi_{(m)}, m = 1, \dots, M$ sind so gewählt, dass folgende Funktion bzgl. der gewählten Gewichte maximiert wird:

$$\sum_{m=1}^M \sum_{m'=1}^M \text{Var} [\xi_{(m)}] |\text{cor}(\xi_{(m)}, \xi_{(l)})|. \quad (2.15)$$

Korrelationsgewichtung: Für die Korrelationsgewichtung ist die Optimierungsfunktion die gleiche wie bei der Vorzeichengewichtung, allerdings wird jetzt über die quadrierten Korrelationskoeffizienten maximiert:

$$\sum_{m=1}^M \sum_{m'=1}^M \text{Var} [\xi_{(m)}] (\text{cor}(\xi_{(m)}, \xi_{(l)}))^2. \quad (2.16)$$

Der Unterschied besteht also lediglich im Betrag bzw. im Quadrat der Korrelationsgewichte. Das führt allerdings dazu, dass bei der Vorzeichengewichtung vor allem Wert darauf gelegt wird, dass die entsprechende LV $\xi_{(m)}$ zu möglichst vielen Nachbar LV hohe Korrelationen aufweist, während bei der Korrelationsgewichtung die Betonung darauf liegt, möglichst hohe Korrelationen auszuweisen, auch wenn diese nur auf relativ wenige Nachbarn zutrifft. Gleichzeitig werden durch die Varianzkomponente $\text{Var} [\xi_{(m)}]$ solche LV bevorzugt, die sich gut in den Beziehungen zu den entsprechenden MV widerspiegeln. In diesem Sinne wird ein

Kompromiss zwischen dem Hauptkomponentenansatz der PCA, möglichst hohe Erklärung der MV, und dem Ansatz der Kanonischen Korrelation, möglichst hohe Korrelation der latenten Variablen untereinander, erzielt.

Modus B: Vorzeichengewichtung: Der PLS Algorithmus liefert ein Maximum für die Summe der Beträge der Korrelationskoeffizienten der LV im Strukturmodell. Die LV $\xi_{(m)}$, $m = 1, \dots, M$ sind so gewählt, dass folgende Funktion bzgl. der gewählten Gewichte maximiert wird:

$$\sum_{m=1}^M \sum_{m'=1}^M |\text{cor}(\xi_{(m)}, \xi_{(l)})|. \quad (2.17)$$

Korrelationsgewichtung: Für die Korrelationsgewichtung ist die Optimierungsfunktion die gleiche wie bei der Vorzeichengewichtung, allerdings wird jetzt wieder über die quadrierten Korrelationskoeffizienten maximiert:

$$\sum_{m=1}^M \sum_{m'=1}^M (\text{cor}(\xi_{(m)}, \xi_{(l)}))^2. \quad (2.18)$$

Der Modus B berücksichtigt also nur die Zusammenhänge im Strukturmodell, gemessen in Form von Korrelationskoeffizienten. Er entspricht somit dem Vorgehen in der Kanonischen Korrelationsanalyse mit unterschiedlichen Optimierungsvorschriften. Es werden lediglich die Beziehungen im Strukturmodell optimiert. Erklärungseigenschaften bzgl. der MV werden nicht berücksichtigt. Im Kontext des formativen Messmodells werden Fehlerterme vor allen Dingen in die MV-Blöcke verschoben.

Der Modus A hingegen versucht sowohl den Erklärungsanteil im Strukturmodell als auch im Messmodell zu berücksichtigen. Im Sinne des reflektiven Messmodells werden LV gebildet, die Fehlerterme im Strukturmodell und im Messmodell klein halten. Die Varianzkomponente im Optimierungsziel von Modus A (2.15, 2.16) ist im Wesentlichen die Kovarianzmatrix der zugehörigen MV

$$\begin{aligned} \text{Var} [\xi_{(m)}] &= \text{Var} [X_{(m)}\omega_{(m)}] \\ &= \omega'_{(m)} X'_{(m)} X_{(m)} \omega_{(m)} = \omega'_{(m)} \text{Cov} [X_{(m)}] \omega_{(m)}. \end{aligned}$$

Die Kovarianzen in den einzelnen MV-Blöcken spielen also einen entscheidenden Unterschied zwischen den Modi A und B. Sind die MV in den einzelnen Blöcken unabhängig, liefern Modus A und B, bis auf Standardisierungen der MV, dieselben Ergebnisse.

Die für beide Modi ableitbaren Optimierungseigenschaften hängen von den in (PLS1) ermittelten Zusammenhangskoeffizienten ab. In diesem Zusammenhang spielt das Konvergenzverhalten eine Rolle. Dieses ist noch nicht theoretisch bewiesen, alle praktischen Anwendungen zeigen aber, dass scheinbar immer Konvergenz eintritt.

Die angeführten Optimierungskriterien sind zunächst nur heuristisch erklärt und ein exakter Beweis steht noch aus. In MATHES (1993b) wurde gezeigt, dass bei Annahme der Konvergenz des PLS-Algorithmus zumindest für Modus B die PLS-Lösungen stationäre Punkte der obigen Maximierungsprobleme (2.17, 2.18) sind. Ob sie wirklich die globalen Maxima sind ist bisher nur Vermutung, allerdings durch eine große Zahl praktischer Beispiele erhärtet. Es existiert ein Arbeitspapier von Tenenhaus (2004), welches für den Modus A und Vorzeichengewichtung (2.15) ebenfalls zeigt, dass ein stationärer Punkt erreicht wird. Die in BETZIN (2000) angeführten Überlegungen zur Gestalt der PLS-Lösungen als verallgemeinerte Eigenwertprobleme nach Konvergenzeintritt liefern dieselben Ergebnisse für beide Modi und jeweils beide Gewichtungsarten.

Berechnung der LV-Scores

Schließlich werden im abschließenden Iterationsschritt die jeweils neuen LV berechnet.

(PLS3) LV Berechnung:

$$\xi_{(m)_{(k+1)}} = X_{(m)} \omega_{(m)_{(k+1)}} c_{(m)_{(k+1)}}. \quad (2.19)$$

Die $\xi_{(m)}$ bilden damit eine gewichtete Summe der zugehörigen MV mit den unter (2.13) bzw. (2.14) berechneten Gewichten $\omega_{(m)}$. Die Größe $c_{(m)}$ dient wiederum zur Normierung.

Die Iteration wird nach Erreichen einer Konvergenzschranke (Stabilitätstest)

$$\sum_{m=1}^M \|\xi_{(m)_{(k+1)}} - \xi_{(m)_{(k)}}\|^2 < \epsilon,$$

mit einem vorher festgelegten „kleinen“ ϵ , abgebrochen.

Wie im ALS-Algorithmus zur Bestimmung der ersten Hauptkomponente eines MV-Blockes werden im PLS-Algorithmus wechselseitig LV $\xi_{(m)}$ und Gewichte $\omega_{(m)}$ innerhalb der MV-Blöcke bestimmt. Der Schritt (PLS3) ist dann hier das Pendant zu (ALS2) und (PLS2) entspricht dem Schritt (ALS1). Zusätzlich müssen hier allerdings die Beziehungen des Strukturmodells berücksichtigt werden, was mit dem Schritt (PLS1) geschieht.

3 Die Lösung des Strukturgleichungsmodells

Die bereits in der Einführung angesprochene zweite Stufe des PLS-Algorithmus kann jetzt mit Hilfe der geschätzten Werte der LV in Angriff genommen werden.

Die Modellparameter des äußeren Mess- und inneren Strukturmodells ergeben sich, nachdem die LV $\xi_{(m)}$ geschätzt wurden, als Lösungen von Kleinst-Quadrat-Regressionen.

MGS-Lösung: Für das reflektive Messmodell ergeben sich die Ladungskoeffizienten $\lambda_{(m)}$ gemäß (1.4) zu

$$(\xi'_{(m)}\xi_{(m)})^{-1}\xi'_{(m)}X_{(m)} =: \lambda_{(m)}.$$

Der in einer allgemeinen OLS-Lösung auftretende Koeffizient $(\xi'_{(m)}\xi_{(m)})^{-1}$ entfällt bei der Berechnung der eigentlichen Lösung, da die berechneten LV normiert sind und damit der Koeffizient ebenfalls 1 wird.

Für das formative Messmodell (1.6) sind die Gewichtskoeffizienten gerade die Gewichtskoeffizienten aus dem letzten Iterationszyklus gemäß (2.7) mit der entsprechenden Normierung.

Bezeichnet man mit $\xi_{C_m^P}$ den Vektor der Vorgänger von $\xi_{(m)}$ und mit $\gamma_{C_m^P}$ den Vektor der entsprechenden Regressionskoeffizienten in (1.1), so ergibt sich folgende

SGS-Lösung:

$$(\xi'_{C_m^P}\xi_{C_m^P})^{-1}\xi'_{C_m^P}\xi_{(m)} = \gamma_{C_m^P}.$$

$\gamma_{C_m^P}$ sind dann wieder die in (1.1) auftretenden Koeffizienten der entsprechenden Regressionsgleichungen des Strukturmodells.

Die Interpretation und Bewertung der Ladungen, Gewichte und Regressionskoeffizienten erfolgt dann entsprechend den zugehörigen Regressionsgleichungen bzw. Faktorgleichungen. Die Validierung von PLS-Pfadmodellen ist Thema des nachfolgenden Beitrags.