Einführung in die PLS-Pfadmodellierung

Dipl.-Wirtsch.-Ing. Jörg Henseler, Kaiserslautern

Strukturgleichungsmodelle mit latenten Variablen erfreuen sich großer Beliebtheit in der empirischen Wirtschafts- und Sozialforschung. Das bislang dominierende Verfahren, die Kovarianzstrukturanalyse, stellt Anforderungen an das Datenmaterial, die in der Forschungsrealität oft nicht gegeben sind. Die PLS-Pfadmodellierung ist ein alternatives Verfahren, das geringere Verteilungsanforderungen stellt. Dieser Beitrag erläutert überblicksartig den Aufbau, die Schätzung und die Interpretation von PLS-Pfadmodellen.

Dipl.-Wirtsch.-Ing. Jörg Henseler ist Assistant Professor für Marketing an der Nijmegen School of Management, Radboud Universiteit Nijmegen (Niederlande). Bevorzugte Forschungsgebiete: Strukturgleichungsmodellierung, Dienstleistungsmarketing, Beziehungsmarketing.

1. Einsatz der PLS-Pfadmodellierung

Zur empirischen Überprüfung komplexer Wirkzusammenhänge erfreut sich die in den Sozialwissenschaften entwickelte **Strukturgleichungsmodellierung** mit **latenten Variablen** großer Beliebtheit. Ein Hauptgrund dafür ist sicherlich, dass sie simultan einerseits die Qualität der Konstruktmessung hinsichtlich der Reliabilität und Validität untersucht und andererseits die Stärke der Beziehungen zwischen den Konstrukten schätzt. Bei der Strukturgleichungsmodellierung mit latenten Variablen konkurrieren insbesondere zwei Verfahren miteinander: Auf der einen Seite die **Kovarianzstrukturanalyse**, im deutschen Sprachraum oft auch als **Kausalanalyse** bezeichnet, die maßgeblich von *Jöreskog* und *Sörbom* (siehe z. B. 1979) vorangebracht und als LISREL bezeichnet wurde, und auf der anderen Seite die von *Wold* (1966) entwickelte Partial-Least-Squares-Pfadmodellierung (PLS). *Fornell* (1982) zählt beide Verfahren zu einer „zweiten Generation multivariater Analysemethoden".

Die PLS-Pfadmodellierung wurde vorwiegend für die Analyse von Daten entwickelt, für die nur geringe theoretische Erklärungsansätze bestehen. Anstatt – wie dies bei LISREL geschieht – die Kovarianzmatrix zu untersuchen, wählte *Wold* eine Analyse der Rohdatenmatrix. Dies erlaubt, neben der Ermittlung der Stärke der Wirkzusammenhänge, die explizite Schätzung der Werte der latenten Variablen. Da die Schätzung von PLS-Pfadmodellen – wie noch gezeigt wird – lediglich auf Kleinste-Quadrat-Schätzungen beruht, kommt das Verfahren ohne strikte Verteilungsannahmen aus. Ein weiterer Vorteil der PLS-Pfadmodellierung im Vergleich zur Kovarianzstrukturanalyse ist eine geringere erforderliche Stichprobengröße. Die für die Modellschätzung mit PLS erforderliche Stichprobengröße bemisst sich nach der in Hinblick auf die Anzahl der Regressoren komplexesten im Modell anzutreffenden multiplen Regression; *Barclay/Higgins/Thompson* (1995, S. 292) fordern als Daumenregel zehn Datensätze mal die Anzahl der Regressoren der komplexesten Regression.

Die Kovarianzstrukturanalyse kommt zwar in der aktuellen Literatur häufiger zum Einsatz, Meta-Untersuchungen, wie z. B. die Arbeit von *Hulland* (1999), belegen jedoch, dass sich die PLS-Pfadmodellierung einer wachsenden Beliebtheit erfreut. Während es für die Kovarianzstrukturanalyse im deutschsprachigen Raum viele einführende Schriften gibt (vgl. z. B. *Homburg*, 1992; *Hildebrandt/Homburg*, 1998), lässt sich keine vergleichbare Einführung für die PLS-Pfadanalyse finden. Diese Lücke möchte der vorliegende Beitrag füllen. Hierbei wird zunächst der Aufbau von PLS-Pfadmodellen beschrieben, bevor dann der eigentliche Schätzalgorithmus vorgestellt wird. Abschließend wird auf die Interpretation von PLS-Pfadmodellen eingegangen.

2. Aufbau von PLS-Pfadmodellen

Strukturgleichungsmodelle mit latenten Variablen, also auch PLS-Pfadmodelle, bestehen typischerweise aus zwei Submodellen: dem Messmodell und dem Strukturmodell. *Abb. 1* veranschaulicht dies. Die Spezifikation der beiden Submodelle bildet die Grundlage eines PLS-Pfadmodells; man spricht hier auch von der **Prädiktorspezifikation**. Sie ist eine wichtige durch den Forscher zu erbringende Vorleistung.

2.1. Das Messmodell

Das **Messmodell** (oder **äußere Modell**) gibt an, wie die latenten Variablen ξ_j mit den manifesten Variablen x_{jh} in Verbindung gebracht werden. Bei einer latenten Variablen „handelt es sich .. um eine nicht direkt messbare Größe" (*Homburg/Giering*, 1996, S. 6), die erst durch den Zusammenhang mit direkt greifbaren, d.h. manifesten Variablen, messbar wird. Bei der Spezifikation des Messmodells müssen zum einen den latenten Variablen manifeste Variablen zugeordnet werden, die jene möglichst gut messen. Zum anderen ist die Form des Messmodells festzulegen, und zwar reflektiv oder formativ. Eine ausführliche Diskussion zum Einsatz formativer und reflektiver Messmodelle findet sich bei *Eggert/Fassott* (2003) sowie *Jarvis/MacKenzie/Podsakoff* (2003).

Bei der Verwendung des **reflektiven Messmodells** unterstellt man, dass die latente Variable die zugehörigen manifesten Variablen verursacht. Die Pfeile in der *Abb. 2* symbolisieren die unterstellte Wirkrichtung. Jede manifeste

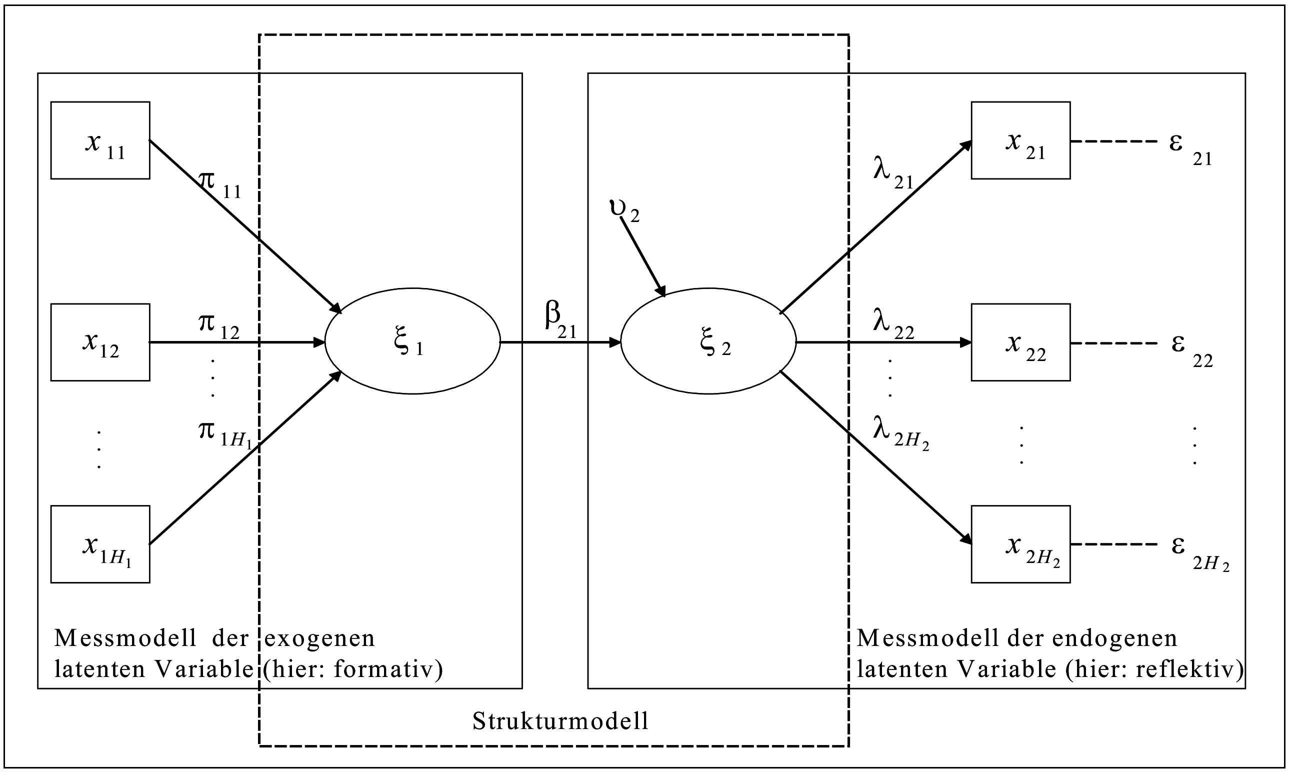

Abb. 1: Modell mit zwei latenten Variablen

Variable ist mit ihrer latenten Variable mittels einfacher Regression verbunden (*Tab. 1* auf S. 74 enthält ein Variablenverzeichnis):

$$x_{jh} = \lambda_{jh}\,\xi_j + \varepsilon_{jh} \tag{1}$$

Das **formative Messmodell** spiegelt einen umgekehrten Wirkzusammenhang wider: Hier verursachen die Werte der manifesten Variablen die Ausprägung der zugeordneten latenten Variable. Die Richtung der Pfeile in *Abb. 3* verdeutlicht dies. Ändert sich einer der Indikatoren, so verändert sich notwendigerweise auch der Wert der latenten Variable. Mathematisch betrachtet ist die latente Variable ξ_j eine Linearkombination ihrer manifesten Variablen x_{jh} (vgl. *Bollen/Lennox*, 1991, S. 306). Die Koeffizienten π_{jh} geben dabei die Gewichtung der Indikatoren x_{jh} bei ihrer linearkombinatorischen Verrechnung zu der latenten Variable ξ_j an. Wird das Messmodell als fehlerbehaftet angesehen, so bezeichnet δ_j den Fehlerterm der Messung (vgl. *Eggert/Fassott*, 2003, S. 2):

$$\xi_j = \sum_h \pi_{jh}\,x_{jh} + \delta_j \tag{2}$$

2.2. Das Strukturmodell

Das **Strukturmodell** (oder **innere Modell**) beschreibt die Beziehungen zwischen den einzelnen latenten Variablen. Im Strukturmodell lassen sich latente Variablen in Abhängigkeit der übrigen latenten Variablen darstellen:

$$\xi_j = \sum_i \beta_{ji}\,\xi_i + \upsilon_j \tag{3}$$

Eine latente Variable, die von keiner anderen latenten Variablen abhängt, wird als **exogen** bezeichnet. Im anderen Fall spricht man von **endogen**.

Die Spezifikation des Strukturmodells liegt ebenso wie die des Messmodells in den Händen des Forschers. Beim Strukturmodell sollte er solche latente Variablen mit Pfeilen verbinden, zwischen denen er – ausgehend z. B. von bestehenden Theorien oder qualitativen Voruntersuchungen – einen Wirkzusammenhang vermutet. Eine notwendige Bedingung an ein gültiges Strukturmodell ist die sog. **Rekursivität**: Innerhalb des Strukturmodells darf es keine kausale Schleife geben, d.h. es darf von keiner latenten Variablen eine Kette von Pfeilen ausgehen, die direkt oder indirekt über andere latente Variablen wieder auf die ursprüngliche latente Variable zeigt.

3. Das Schätzverfahren für PLS-Pfadmodelle

Das Schätzverfahren setzt sich zusammen aus einer Initialisierung, dem Algorithmus zur Schätzung der latenten Variablen sowie der abschließenden Ermittlung der Strukturgleichungsparameter. Zunächst soll die Funktionsweise des Schätzverfahrens bei einem einfachen Modell mit nur zwei latenten Variablen aufgezeigt werden, bevor im darauffolgenden Abschnitt das Verfahren für allgemeine Modelle mit mehreren latenten Variablen erläutert wird.

3.1. Schätzung eines Modells mit zwei latenten Variablen

Die Funktionsweise des Schätzverfahrens lässt sich am besten anhand eines Modells mit zwei Konstrukten demonstrieren. Für das Modell aus *Abb. 1* würde das Schätzverfahren wie folgt ablaufen (vgl. *Barclay/Higgins/Thompson*, 1995, S. 292):

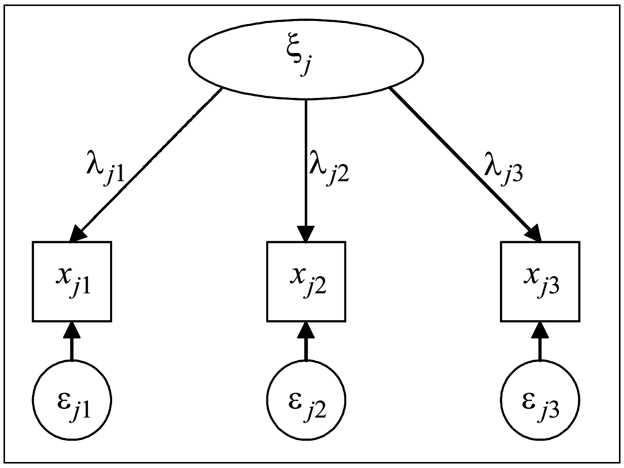

Abb. 2: Reflektives Messmodell der latenten Variablen ξ_j mit drei manifesten Variablen x_{jh} und den zugehörigen Ladungen λ_{jh} und Messfehlern ε_{jh}

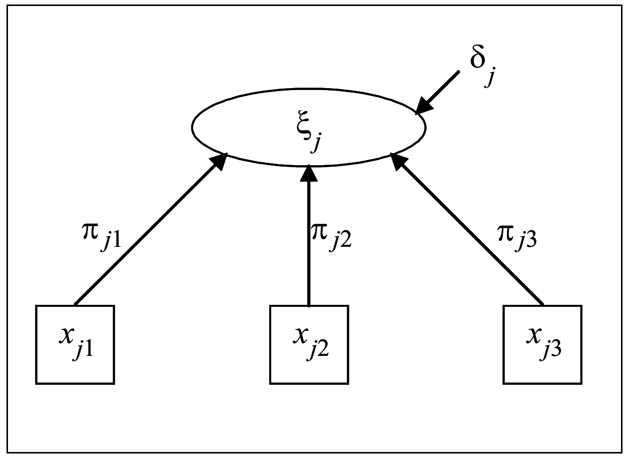

Abb. 3: Formatives Messmodell der latenten Variablen ξ_j mit drei manifesten Variablen x_{jh} sowie den zugehörigen Gewichten π_{jh}

Schritt 0: Für zumindest eine latente Variable ist ein Startwert erforderlich. Ohne Beschränkung der Allgemeinheit wird in der ersten PLS-Iteration ξ_2 gleich einer beliebigen zugehörigen manifesten Variable gesetzt, d.h. z. B. $\xi_2 = x_{21}$.

Schritt 1: Mittels einer multiplen Kleinste-Quadrate-Regression mit ξ_2 als abhängiger Variable und $x_{11}, ..., x_{1H1}$ als unabhängigen Variablen werden die Regressionsgewichte $\pi_{11}, ..., \pi_{1H1}$ geschätzt. Im Prinzip wird so getan, als gelte $\beta_{21} = 1$.

Schritt 2: Nun lässt sich ξ_1 bilden als Linearkombination aus $x_{11}, ..., x_{1H1}$ mit den Gewichten $\pi_{11}, ..., \pi_{1H1}$, d.h. $\xi_1 = \sum_{h=1}^{H_1} \pi_{1h} \cdot x_{1h}$.

Schritt 3: Die Ladungen $\lambda_{21}, ..., \lambda_{2H2}$ werden dann mittels einer Serie einfacher Regressionen mit ξ_1 als unabhängiger Variable und $x_{21}, ..., x_{2H2}$ als abhängigen Variablen geschätzt. Die Ladungen $\lambda_{21}, ..., \lambda_{2H2}$ werden in Gewichte umgeformt.

Schritt 4: Mit diesen Gewichten wird ein neuer Schätzwert für ξ_2 als Linearkombination aus $x_{21}, ..., x_{2H2}$ ermittelt.

Die Schritte 1 bis 4 werden so lange wiederholt, bis ein vorzugebendes Abbruchkriterium erfüllt ist. Die Beschreibung des Algorithmus lässt seine Funktionsweise leicht erkennen: Während man einen Teil des Modells als gegeben betrachtet, wird der andere Teil mittels Regressionen neu berechnet. Daher stammt auch die Bezeichnung Partial Least Squares (PLS), also „partielle kleinste Quadrate".

Das Messmodell inklusive der Werte für ξ_1 und ξ_2 ist nun bestimmt; es fehlt noch die Schätzung des Strukturmodells. Den gesuchten Schätzer β_{21} liefert eine einfache Regression mit, nun bekannten, ξ_1 als unabhängiger und ξ_2 als abhängiger Variable. Aus dieser Regression resultiert ein R^2 als Gütemaß dafür, wie gut ξ_z von ξ_1 erklärt wird.

3.2. Schätzung eines allgemeinen Pfadmodells mit mehr als zwei latenten Variablen

Bei einem allgemeinen Modell mit mehr als zwei Konstrukten, wie es *Abb. 4* zeigt, muss der Algorithmus modifiziert werden. Aus Gründen der Übersichtlichkeit wird für die folgende Darstellung des Algorithmus die Annahme getroffen, dass alle manifesten Variablen standardisiert sind, d.h. einen Mittelwert von 0 und eine Standardabweichung von 1 besitzen. Eine Beschreibung des Algorithmus ohne diese Annahme findet sich z. B. bei *Chatelin/Esposito Vinzi/Tenenhaus* (2002, S. 8–10).

Um den Algorithmus für mehr als zwei latente Variablen zu erweitern, nutzt man einen Trick: Für die latenten Variablen ξ_j werden zwei unterschiedliche Schätzungen ermittelt: ein äußerer Schätzwert Y_j aus dem Messmodell und ein innerer Schätzwert Z_j aus dem Strukturmodell. Der Algorithmus zur Schätzung der latenten Variablen besteht aus der Initialisierung und zwei jeweils zweiteiligen Schritten. Iterativ wird in jedem Schritt ein Schätzwert festgehalten (abwechselnd der innere und der äußere), während (a) Gewichtungsfaktoren und (b) die jeweils anderen Schätzwerte berechnet werden. Die Schritte des Schätzalgorithmus für die latenten Variablen lauten wie folgt:

Schritt 0: Bei der **Initialisierung** wird ein erster äußerer Schätzwert Y_j für jede latente Variable ξ_j ermittelt. Im Prinzip könnten die Y_j beliebige nichttriviale Linearkombinationen der zugehörigen manifesten Variablen sein. Aus praktischen Erwägungen schlagen *Chatelin/Esposito Vinzi/Tenenhaus* (2002, S. 9f.) als eine Möglichkeit vor, das Gewicht des ersten Indikators einer jeden latenten Variablen auf 1 und die Gewichte der übrigen Indikatoren auf 0 zu setzen. Y_j wird somit gleich der ersten manifesten Variablen gesetzt:

$$\forall\, j: Y_j := \mathrm{x}_{j1} \qquad (4)$$

Schritt 1a: Zur Vorbereitung der inneren Schätzung der latenten Variablen dient die **Schätzung** der **inneren Gewichte** e_{ji} . Hierfür existieren drei mögliche Schemata:

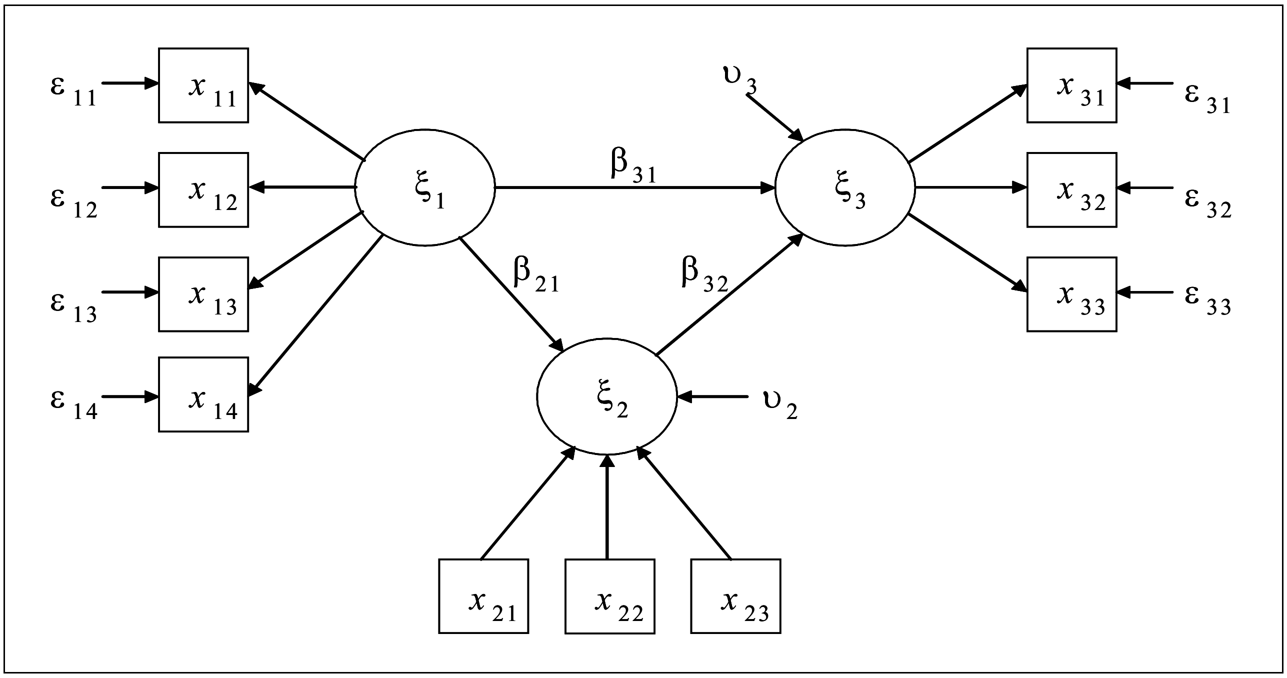

Abb. 4: Beispiel für ein Modell mit drei latenten Variablen

- Sofern das PLS-Pfadmodell Pfeile zwischen den latenten Variablen ξ_i und ξ_j aufweist, werden beim **Zentroidschema** die inneren Gewichte e_{ji} gleich dem Vorzeichen der Korrelation zwischen Y_i und Y_j gesetzt, ansonsten erhalten sie den Wert 0:

$$\forall\, j,i: e_{ji} := \begin{cases} \text{sgn}(\text{korr}(Y_i, Y_j)), & \exists \text{Pfeilbeziehung zwischen } \xi_i \text{ und } \xi_j; \\ 0, & \text{sonst.} \end{cases} \quad (5)$$

Das Zentroidschema besitzt allerdings bei betragsmäßig sehr kleinen Korrelationen den Nachteil, dass es sich sprunghaft verändert.

- Diesen Nachteil behebt das **Faktorgewichtungsschema**. Beim Faktorgewichtungsschema wird nicht das Vorzeichen der Korrelation, sondern die Korrelation selbst verwendet:

$$\forall\, j,i: e_{ji} := \begin{cases} \text{korr}(Y_i, Y_j), & \exists \text{Pfeilbeziehung zwischen } \xi_i \text{ und } \xi_j; \\ 0, & \text{sonst.} \end{cases} \quad (6)$$

- Eine Weiterentwicklung des Faktorgewichtungsschemas ist das **Pfadgewichtungsschema**. Für jede latente Variable ξ_j wird unterschieden zwischen **Vorgängern**, d.h. latenten Variablen, von denen aus ein Pfeil auf ξ_j zeigt, und **Nachfolgern**, d.h. latenten Variablen, auf die ein von ξ_j ausgehender Pfeil zeigt. Für Nachfolger wird weiterhin das Faktorgewichtungsschema verwendet. Für Vorgänger ξ_i werden die inneren Gewichte gleich den Regressionskoeffizienten b_{ji} der multiplen Regression mit allen Vorgängern als unabhängigen Variablen und ξ_j als abhängiger Variable gesetzt. Die Regressionskoeffizienten b_{ji} sind zugleich provisorische Pfadwerte des Strukturmodells.

$$\forall\, j,i: e_{ji} := \begin{cases} b_{ji}, & \exists \text{Pfeilbeziehung von } \xi_i \text{ nach } \xi_j; \\ \text{korr}(Y_i, Y_j), & \exists \text{Pfeilbeziehung von } \xi_j \text{ nach } \xi_i; \\ 0, & \text{sonst.} \end{cases} \quad (7)$$

Schritt 1b: Bei der **inneren Schätzung** der **latenten Variablen** erhält man für die latente Variable ξ_j einen Schätzwert Z_j als gewichtete Summe der übrigen latenten Variablen. Die Gewichte e_{ji} wurden in Schritt 1a ermittelt. Die Faktoren φ_j dienen zur Standardisierung der Z_j.

$$Z_j := \varphi_j \cdot \sum_i e_{ji}\, Y_i \quad (8)$$

Schritt 2a: Die **Schätzung** der **äußeren Gewichte** hängt maßgeblich von der Wahl des Messmodells für jede einzelne latente Variable ab. Für reflektive und formative Messmodelle hat sich jeweils eine Methode etabliert. Bei reflektiven Messmodellen kommt der sog. **Modus A** zum Einsatz. Beim Modus A entspricht das äußere Gewicht π_{jh} dem Regressionskoeffizienten der einfachen Regression mit der manifesten Variablen x_{jh} als abhängiger und der inneren Schätzgröße Z_j als unabhängiger Variable:

$$\pi_{jh} := cov\,(x_{jh}, Z_j) \quad (9)$$

Da sowohl die manifesten Variablen x_{jh} als auch die inneren Schätzgrößen Z_j standardisiert sind, entspricht die Kovarianz dem besagten Regressionskoeffizienten. Auf diesem Weg erhalten diejenigen manifesten Variablen, die sich einen größeren Teil der Varianz mit der inneren Schätzgröße der latenten Variable teilen, erwartungsgemäß ein größeres Gewicht. Bei formativen Messmodellen dominiert der sog. **Modus B**. Beim Modus B ergibt sich der Vektor π_j der Gewichte π_{jh} als Regressionskoeffizientenvektor aus der multiplen Regression mit der inneren

Schätzgröße Z_j als abhängiger Variable und den zur latenten Variablen ξ_j gehörenden manifesten Variablen x_{jh} als unabhängigen Variablen. Reiht man die Spaltenvektoren der manifesten Variablen x_{jh} aneinander zu einer Matrix X_j, so lässt sich der Regressionskoeffizientenvektor mittels der Matrixform der multiplen Regression darstellen:

$$\pi_j := (X_j^T X_j)^{-1} X_j^T Z_j \tag{10}$$

Schritt 2b: Bei der **äußeren Schätzung** der **latenten Variablen** werden alle latenten Variablen anhand der ihnen zugeordneten manifesten Variablen geschätzt. Für die latenten Variablen ξ_j ermittelt man äußere Schätzwerte Y_j als Linearkombinationen der manifesten Variablen (die Faktoren f_j dienen zur Standardisierung der Y_j):

$$Y_j := f_j \sum \pi_{jh}\, x_{jh} \tag{11}$$

Die Schritte 1a bis 2b werden so lange wiederholt, bis ein vorzugebendes Abbruchkriterium erfüllt ist. Als Abbruchkriterium dient üblicherweise die Summe der Änderung der Gewichte von einer Iteration zur nächsten. Sobald die Summe der quadrierten Differenzen der Gewichtsänderungen einen bestimmten Wert – *Wold* (1982, S. 14) schlägt z. B. den Wert 10^{-5} vor – unterschreitet, wird der Algorithmus abgebrochen. Der Algorithmus zur Schätzung der latenten Variablen endet stets mit der äußeren Schätzung der latenten Variablen. Aus den manifesten Variablen und den äußeren Gewichten kann man die Werte der latenten Variablen auch außerhalb einer PLS-Software berechnen.

Während das Messmodell nun bestimmt ist, fehlt noch die Schätzung des Strukturmodells. Dies geschieht mittels multipler Regressionen: Für jede endogene latente Variable ξ_j werden die Pfadwerte β_{ji} als Regressionskoeffizienten aus der Regression mit ξ_j als abhängiger und all ihren Vorgängern ξ_i als unabhängigen Variablen berechnet.

4. Interpretation von PLS-Pfadmodellen

Der PLS-Algorithmus zur Schätzung der latenten Variablen liefert Schätzer für die Gewichte bzw. Ladungen der manifesten Variablen, die Werte der latenten Variablen sowie die Pfadkoeffizienten zwischen den latenten Variablen. Bei formativen Messmodellen besagen die Gewichte der manifesten Variablen, wie stark der Einfluss der jeweiligen manifesten Variable auf die zugehörige latente Variable ist. Bei reflektiven Messmodellen drücken die Ladungen der manifesten Variablen aus, wie groß die gemeinsame Varianz mit der latenten Variablen ist, d.h. wie gut sie diese reflektieren. Bei der Interpretation der Parameterwerte sowie der latenten Variablen ist zu beachten, dass die latenten Variablen bezüglich des Vorzeichens nicht bestimmt sind. Ist das Vorzeichen der Gewichte bzw. Ladungen derjenigen manifesten Variablen negativ, die erwartungsgemäß positiv mit der latenten Variablen korrelieren sollten, so muss die latente Variable in das logische Gegenteil umgedeutet werden. Misst z. B. eine Reihe von Indikatoren unterschiedliche Facetten von „Loyalität“ und sind ihre Gewichte bzw. Ladungen negativ, so entspricht die zugehörige latente Variable inhaltlich der „Illoyalität“. Dies wirkt sich auf das Strukturmodell aus: Die Pfadwerte als Maß für die Wirkbeziehungen sind unter Beachtung der Orientierung der latenten Variablen zu interpretieren.

Ausgangspunkt für die **Evaluierung** von PLS-Pfadmodellen bilden die R^2 aus den Regressionen der einzelnen endogenen latenten Variablen des Strukturmodells. Sie lassen sich genauso interpretieren wie bei gewöhnlichen Regressionen. Gleiches gilt für die Pfadwerte: Auch sie können als Koeffizienten einer gewöhnlichen (multiplen) Regression angesehen werden – schließlich wurden sie auch so ermittelt.

Die Analyse der R^2 -Änderung gibt Aufschluss darüber, ob eine unabhängige latente Variable einen substanziellen Einfluss auf eine abhängige Variable ausübt. *Chin* (1998, S. 316) betrachtet hierzu die Effektstärke f^2 , die sich aus den unterschiedlichen R^2 berechnet, wenn die betreffende unabhängige latente Variable in die abhängige latente Variable einfließt (R^2_{inkl}) oder nicht (R^2_{exkl}) :

$$f^2 := \frac{R^2_{inkl} - R^2_{exkl}}{1 - R^2_{inkl}}\,. \tag{12}$$

Werte für f^2 von über 0,02 / 0,15 / 0,35 besagen, dass die betreffende unabhängige latente Variable einen kleinen / mittleren / großen Einfluss auf die abhängige latente Variable hat.

Zur **Validierung** der ermittelten Schätzwerte gelangen **Resamplingtechniken** zum Einsatz. Sie erlauben eine Bewertung der Stabilität der ermittelten Modellparameter. Aus der Rohdatenmatrix wird n mal eine Anzahl k von Beobachtungen (mit oder ohne Zurücklegen) gezogen, die jeweils zu einer modifizierten Rohdatenmatrix umgeformt werden. Für jede der n modifizierten Rohdatenmatrizen werden die Modellparameter neu berechnet. Über alle n geschätzten Modelle hinweg wird für jeden Modellparameter Mittelwert und Standardfehler ermittelt. Der Quotient aus Mittelwert und Standardfehler entspricht dem t-Wert, aus dem sich Signifikanzen bzw. Konfidenzintervalle für die Modellparameter ableiten lassen.

In der Modellvalidierung liegt ein Nachteil von PLS-Pfadmodellen: Es existiert – anders als bei der Kovarianzstrukturanalyse – kein globales Gütemaß. Sowohl das Messmodell als auch das Strukturmodell lassen sich nur partiell validieren.

5. Zusammenfassung

Die PLS-Pfadmodellierung ist eine Alternative zur Kovarianzstrukturanalyse. Im Vergleich stellt sie geringere Anforderungen an die statistische Verteilung der Rohdaten sowie die Stichprobengröße und ermittelt explizit Werte für die latenten Variablen.

PLS-Pfadmodelle bestehen aus dem Messmodell und dem Strukturmodell. Mit PLS lassen sich Wirkmodelle sowohl

Variable	Bedeutung der Variable
b_{ji}	vorläufiger Pfadkoeffizient des Pfads von ξ_i nach ξ_j
β_{ji}	Pfadkoeffizient des Pfad von ξ_i nach ξ_j
δ_j	Fehlerterm der latenten Variable ξ_j bei formativem Messmodell
e_{ji}	Gewicht des Pfad von ξ_i nach ξ_j für innere Schätzung
ε_{jh}	Fehlerterm des h-ten Indikators der reflektiv gemessenen latenten Variable ξ_j
φ_j, f_j	Faktoren zur Standardisierung der latenten Variable ξ_j
$h = 1,...,H_j$	Index über die Indikatoren der latenten Variable ξ_j
i, j	Indizes über die latenten Variablen
λ_{jh}	Ladung des h-ten Indikators der latenten Variable ξ_j
π_{jh}	Gewicht des h-ten Indikators der latenten Variable ξ_j
x_{jh}	h-ter Indikator der latenten Variable ξ_j
$X_j = (x_{j1} \quad ... \quad x_{jH_j})$	Matrix aus den Indikatoren der latenten Variable ξ_j
ξ_j	latente Variable
Y_j	äußerer Schätzer für die latente Variable ξ_j
υ_j	Fehlerterm der latenten Variable ξ_j beim Strukturmodell
Z_j	innerer Schätzer für die latente Variable ξ_j

Tab. 1: Variablenverzeichnis

mit reflektiven als auch mit formativen Messmodellen schätzen. Der eigentliche Schätzalgorithmus besteht aus einer Serie von einfachen und multiplen Regressionen: Stets wird ein Teil des Modells festgehalten, während der andere Teil davon ausgehend berechnet wird. Die Parameterwerte von PLS-Pfadmodellen lassen sich analog zu herkömmlichen linearen Regressionen interpretieren. Die Validierung erfolgt mittels Resamplingtechniken. Für PLS-Pfadmodelle stehen jedoch keine globalen Gütemaße zur Verfügung.

Literatur

Barclay, D., Ch. Higgins, R. Thompson, The Partial Least Squares (PLS) Approach to Causal Modeling. Personal Computer Adoption and Use as an Illustration, in: Technology Studies, Vol. 2 (1995), No. 2, S. 285–314.

Bollen, K.A., R. Lennox, Conventional Wisdom on Measurement. A Structural Equation Perspective, in: Psychological Bulletin, Vol. 110 (1991), No. 2, S. 305–314.

Chatelin, Y.M., V. Esposito Vinzi, M. Tenenhaus, State-of-Art on PLS Path Modeling Through the Available Software, HEC Working Paper (2002), No. 764.

Chin, W.W., The Partial Least Squares Approach to Structural Equation Modeling, in: *G.A. Marcoulides* (Ed.), Modern Methods for Business Research, Mahwah, New Jersey 1998, S. 295–337.

Eggert, A., G. Fassott, Zur Verwendung normativer und reflektiver Indikatoren in Strukturgleichungsmodellen. Ergebnisse einer Meta-Analyse und Anwendungsempfehlungen, Kaiserslauterer Schriftenreihe Marketing (2003), Nr. 20.

Fornell, C. (Ed.), A Second Generation of Multivariate Analysis. Methods, Vol. 1, New York 1982.

Hildebrandt, L., Ch. Homburg (Hrsg.), Die Kausalanalyse. Ein Instrument der empirischen betriebswirtschaftlichen Forschung, Stuttgart 1998.

Homburg, Ch., Die Kausalanalyse. Eine Einführung, in: WiSt – Wirtschaftswissenschaftliches Studium, 21. Jg. (1992), S. 499–508.

Homburg, Ch., A. Giering, Konzeptualisierung und Operationalisierung komplexer Konstrukte. Ein Leitfaden für die Marketingforschung, in: Marketing ZFP, 18. Jg. (1996), S. 5–24.

Hulland, J., Use of Partial Least Squares (PLS) in Strategic Management Research. A Review of Four Recent Studies, in: Strategic Management Journal, Vol. 20 (1999), S. 195–204.

Jarvis, C.B., S.B. MacKenzie, P.M. Podsakoff, A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research, in: Journal of Consumer Research, Vol. 30 (2003), No. 2, S. 199–218.

Jöreskog, K.G., D. Sörbom, Advances in Factor Analysis and Structural Equation Models, Cambridge 1979.

Lohmöller, J.B., Latent Variable Path Modeling with Partial Least Squares, Heidelberg 1989.

Wold, H., Non-linear Estimation by Iterative Least Squares Procedures, in: *F.W. David* (Ed.), Research papers in statistics, New York 1966, S. 411–444.

Wold, H., Nonlinear Iterative Partial Least Squares (NIPALS) Modelling. Some Current Developments, in: *P.R. Krishnaiah* (Ed.), Proceedings of the 3rd International Symposium on Multivariate Analysis, Dayton, Ohio 1973, S. 383–407.

Wold, H., Soft Modeling. The Basic Design and Some Extensions, in: *K.G. Jöreskog, H. Wold* (Eds.), Systems under Indirect Observation. Causality, Structure, Prediction. Part I, Amsterdam, New York, Oxford 1982, S. 1–54.