

Prof N. Carlo Lauro

Modelli ad Equazioni Strutturali

slides a cura di
Laura Trincherà Daniela Nappo

Dipartimento di Matematica e Statistica
Università degli Studi di Napoli "Federico II"



Modelli ad Equazioni Strutturali - corso 2009

Lezione 1

Introduzione

I Modelli ad Equazioni Strutturali (Structural Equation Models, SEM)* sono modelli in grado di modellizzare complesse strutture di relazioni di **CAUSALITA'** tra concetti latenti (le *Variabili Latenti, VL*) a partire da un insieme di indicatori reali di solito definite come *Variabili Manifeste (VM)*.

Il termine SEM è da riferirsi a una famiglia molto vasta di metodi statistici in grado di analizzare le relazioni statistiche di causa-effetto e non ad un'unica tecnica.

* Per uno studio approfondito fare riferimento a Bollen 1989, Kaplan 2000 e Tenenhaus et al. 2005

Un po' di storia....chi ha inventato i SEM??

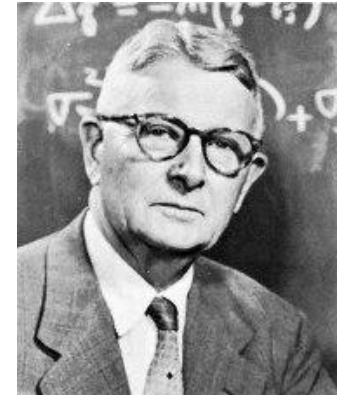
Diverse discipline hanno avuto un ruolo nello sviluppo dei Modelli ad Equazioni Strutturali:

- L'"Econometria" di Sewall Wright e la Path Analysis
- La Psicometria e l'Analisi Fattoriale
- Economisti e Sociologi alla riscoperta dei Modelli Causali
- Otis D. Duncan ed il "ricongiungimento"
- I Modelli ad Equazioni Strutturali e Karl Jöreskog

Sewall Wright e la Path Analysis

Sewall Wright (21 Dicembre 1889 -3 Marzo 1988)

Genetista americano figlio dell'economista Philip Wright



La **Path Analysis** è stata sviluppata negli anni 20 del novecento da S. Wright sia per lo studio dei fenomeni genetici sia come supporto all'attività di ricerca del padre economista.

Lo scopo è quello di studiare le relazioni di dipendenza tra più variabili semplicemente osservando la struttura di correlazione tra le stesse, e cioè la matrice dei coefficienti di correlazione.

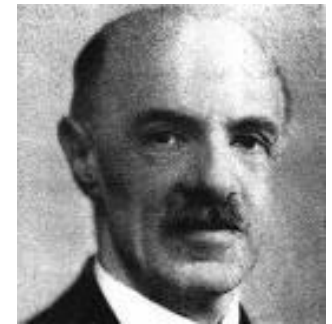
L'innovazione principale della Path Analysis è legata all'introduzione di un nuovo strumento: il **path diagram**

L'Analisi Fattoriale e la Psicometria

Charles Edward Spearman (10 Settembre 1863 - 17 Settembre 1945)

Psicologo inglese

C. Spearman sviluppo le basi dell'**Analisi Fattoriale (AF)** all'inizio del 1900 per misurare l'intelligenza in modo "obbiettivo".



L'idea di base è che l'intelligenza sia un fenomeno **MULTIDIMENSIONALE** e che la correlazione tra più variabili osservate possa essere descritta da un unico "fattore" sottostante

Uno dei maggiori contributi dell'**Analisi Fattoriale** è il concetto di "fattore", in altre parole il concetto di **Variabile Latente**.

Louis Thurstone e l'incontro con Herman Wold

L'idea originale di Spearman è stata modificata nei successivi 40 anni in modo da prendere in considerazione più fattori come "cause" dei legami causali osservati tra più insiemi variabili manifeste.

Louis Thurstone (29 Maggio 1887-30 Settembre 1955)

può essere considerato come il padre dell'Analisi Fattoriale Multipla (da non confondere con quella di Escofier e Pages)

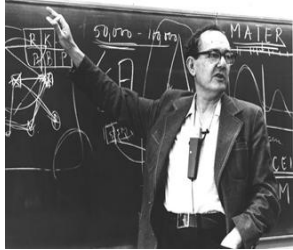


Herman Wold (25 dicembre 1908 - 16 febbraio 1992)

Incontra all'inizio degli anni 50 Thurstone ed insieme organizzano "the Upspsala Symposium on Psychological Factor Analysis". Quest'incontro segna l'inizio dei lavori di H. Wold sui modelli a variabili latenti.



Economisti e Sociologi alla riscoperta dei Modelli Causali



Herbert Simon (15 giugno 1916 -9 febbraio 2001)
Economista - Nobel per l'economia nel 1978

Nel 1954 pubblica un lavoro in cui dimostra che *"sotto certe assunzioni la correlazione è un indice di causalità"*

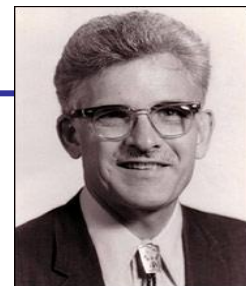
Hubert M. Blalock (23 Agosto 1926 - 8 Febbraio 1991)
Sociologo



Nel 1964 pubblica il libro "Causal Inference in Nonexperimental Research", nel quale definisce dei metodi in grado di effettuare analisi inferenziali causali partendo dalla matrice di correlazione osservata e affronta il problema della "conferma" inferenziale delle relazioni esistenti tra più variabili

Insieme arrivano alla definizione del metodo SIMON-BLALOCK

Economisti e Sociologi alla riscoperta dei Modelli Causali



Otis D. Duncan (2 Dicembre 1921- 16,Novembre 2004)

è stato uno dei più importanti sociologi nel mondo. A lui si deve l'introduzione della Path Analysis di Wright in Sociologia.

Verso la metà degli anni 60 arriva alla conclusione che non c'è differenza tra la Path Analysis di Wright e il modello di di Simon-Blalock.

Successivamente entra in contatto con l'economista Arthur Goldberger ed insieme arrivano alla conclusione che non c'era differenza tra quella che in sociologia era nota come Path Analysis ed i modelli ad equazioni simultanee comunemente usati in econometria.

Insieme organizzano una conferenza nel 1970 a Madison (USA) dove invitano: Karl Jöreskog

I Modelli ad Equazioni Strutturali e K. Jöreskog

Karl Jöreskog è Professore Emerito alla Uppsala University, Svezia.



Alla fine degli anni 50 comincia a collaborare con Herman Wold e discute una tesi sull'Analisi Fattoriale.

Nella seconda metà degli anni 60 entra in contatto con O.D. Duncan e A. Goldberger. Questa collaborazione rappresenta l'incontro tra l'Analisi Fattoriale ed i metodi a variabili latenti e la Path Analysis e i modelli causali.

Nel 1970 Jöreskog presenta ad una conferenza organizzata da Duncan e Goldberger la sua: *Covariance Structure Analysis (CSA) for estimating a linear structural equation system*, più tardi conosciuto come LISREL

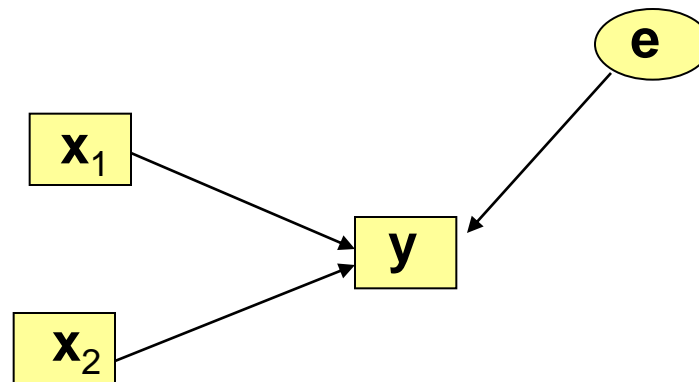
Il Path Diagram

Il Path Diagram è una rappresentazione grafica delle relazioni esistenti tra le variabili sotto esame

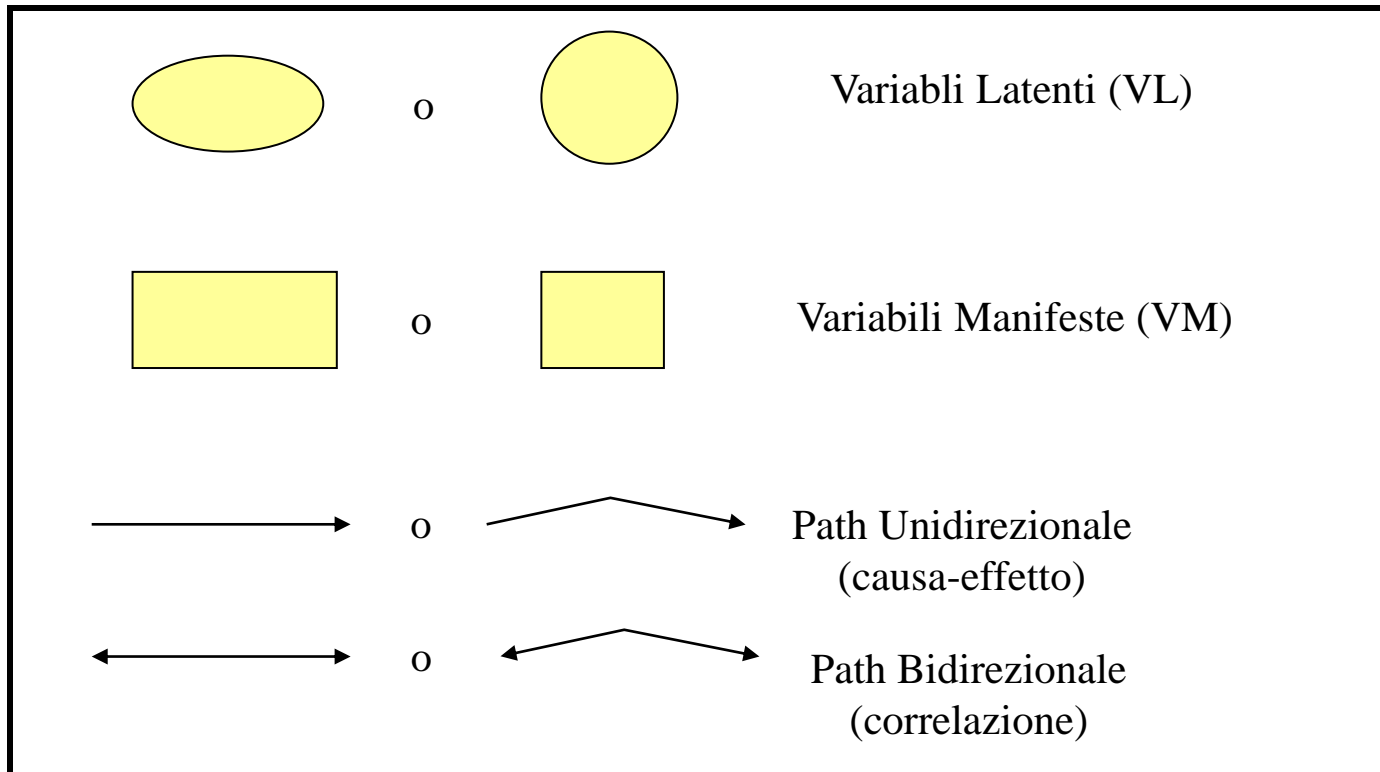
Un semplice modello di regressione multipla del tipo:

$$y = ax_1 + bx_2 + e$$

può essere rappresentato graficamente attraverso un semplice Path Diagram:



Il Path Diagram: i simboli



I Modelli ad Equazioni Strutturali: le notazioni

Le Variabili Latenti sono indicate con lettere greche:

$$\eta = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_m \end{bmatrix}$$

Variabili Latenti ENDOGENE, con
 $m = \#$ VL endogene

$$\xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_j \end{bmatrix}$$

Variabili Latenti ESOGENE con
 $j = \#$ VL endogene

Le Variabili Manifeste sono indicate con lettere romane:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$$

con $p = \#$ VM endogene

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_q \end{bmatrix}$$

con $q = \#$ VM endogene

* Queste notazioni sono quelle presentate da Jöreskog e sono caratteristiche della sua CSA, le notazioni del PLS Path Modeling sono leggermente diverse

L'Analisi Fattoriale ed il concetto di VL

L'obiettivo di una *Analisi Fattoriale (AF)* è quello di spiegare la variabilità esistente tra una serie di variabili manifeste direttamente osservate ed i loro legami in termini di un numero ridotto di variabili latenti: *i* fattori.

In termini più formali:

date *p* variabili manifeste osservate su *n* individui $\mathbf{x}_1 \dots \mathbf{x}_p$ l'obiettivo di una AF è di riformulare le *p* variabili manifeste in termini di *q* fattori "comuni" (con $q < p$) e *p* errori:

$$\begin{aligned} \mathbf{x}_1 &= \lambda_{11}\xi_1 + \lambda_{12}\xi_2 + \dots + \lambda_{1q}\xi_q + \delta_1 \\ \mathbf{x}_2 &= \lambda_{21}\xi_1 + \lambda_{22}\xi_2 + \dots + \lambda_{2q}\xi_q + \delta_2 \\ &\vdots \\ \mathbf{x}_p &= \lambda_{p1}\xi_1 + \lambda_{p2}\xi_2 + \dots + \lambda_{pq}\xi_q + \delta_p \end{aligned}$$

Diagram annotations:

- Factor Loadings**: points to the λ terms in the equations.
- Common factors**: points to the ξ terms in the equations.
- Unique factors**: points to the δ terms in the equations.

L'Analisi Fattoriale Esplorativa e Confermativa

Analisi Fattoriale Esplorativa

ha come obiettivo quello di determinare come e in che modo le variabili manifeste osservate sono legate ad uno o più fattori latenti.

Le relazioni tra le variabili osservate e le variabili latenti sono quindi sconosciute o incerte.

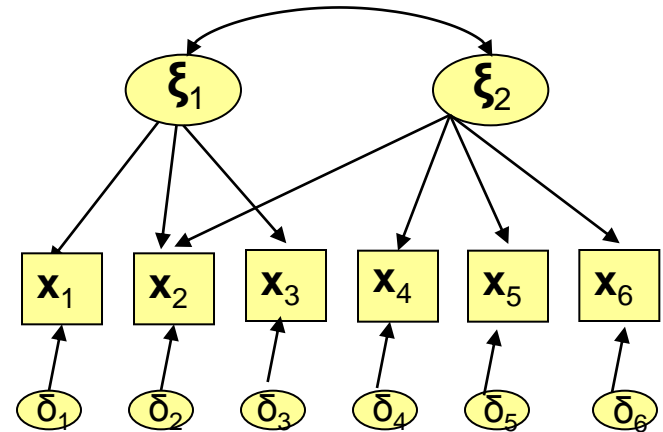
Analisi Fattoriale Confermativa

ha come scopo quello di testare statisticamente le relazioni causali esistenti tra le variabili manifeste e uno o più fattori latenti.

Le relazioni tra le variabili manifeste e le variabili latenti sono quindi note a priori sulla base di teorie o sulla base di esperimenti empirici

L'Analisi Fattoriale Confermativa: un esempio

$$\begin{aligned}x_1 &= \lambda_{11} \xi_1 + \delta_1, \\x_2 &= \lambda_{21} \xi_1 + \lambda_{22} \xi_2 + \delta_2 \\x_3 &= \lambda_{31} \xi_1 + \delta_3 \\x_4 &= \lambda_{42} \xi_2 + \delta_4 \\x_5 &= \lambda_{52} \xi_2 + \delta_5 \\x_6 &= \lambda_{62} \xi_2 + \delta_6\end{aligned}$$



$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_6 \end{bmatrix} \quad \Lambda_x = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & \lambda_{22} \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \end{bmatrix}$$

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \quad \delta = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_6 \end{bmatrix}$$

$$\mathbf{x} = \Lambda_x \xi + \delta$$

L'Analisi Fattoriale Confermativa: un esempio

$$\mathbf{x} = \Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta}$$

Assumendo che:

→ le VM, i fattori comuni ed i fattori specifici abbiano un media pari a zero

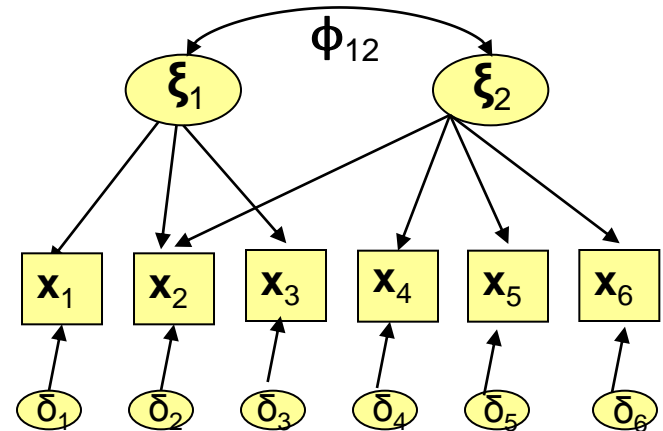
→ che i fattori comuni siano standardizzati

→ i fattori specifici ed i fattori comuni non covarino tra loro

e che: $\Sigma = E(\mathbf{x}\mathbf{x}')$ è la matrice di var/cov delle VM

$\Theta = E(\boldsymbol{\delta}\boldsymbol{\delta}')$ è la matrice di var/cov tra i fattori specifici (o errori di misura)

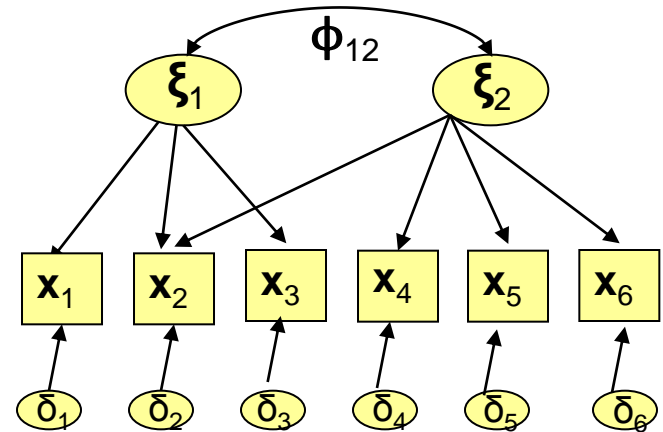
$\Phi = E(\boldsymbol{\xi}\boldsymbol{\xi}')$ è la matrice di var/cov tra i fattori comuni



L'Analisi Fattoriale Confermativa: un esempio

$$\mathbf{x} = \Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta}$$

La matrice di var/cov tra le VM può essere riscritta in termini di parametri del modello, come:



$$\Sigma = E(\mathbf{xx}') = E\left[(\Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta})(\Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta})' \right] =$$

$$= E\left[(\Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta})(\Lambda_x' \boldsymbol{\xi}' + \boldsymbol{\delta}') \right] =$$

$$= \Lambda_x E(\boldsymbol{\xi} \boldsymbol{\xi}') \Lambda_x' + \Lambda_x E(\boldsymbol{\xi} \boldsymbol{\delta}') + E(\boldsymbol{\delta} \boldsymbol{\xi}') \Lambda_x' + E(\boldsymbol{\delta} \boldsymbol{\delta}')$$

$$\Sigma = \Lambda_x \Phi \Lambda_x' + \Theta$$

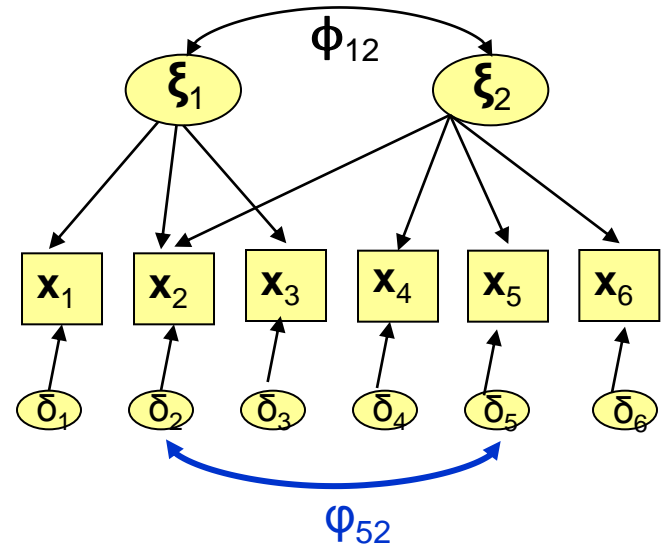
L'Analisi Fattoriale Confermativa: un esempio

$$\Sigma = \Lambda_x \Phi \Lambda'_x + \Theta$$

Dove:

$$\Lambda_x = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & \lambda_{22} \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \end{bmatrix}$$

$$\Phi = \begin{bmatrix} 1 & \phi_{12} \\ & 1 \end{bmatrix}$$



I fattori comuni (le VL) sono standardizzate!

$$\Theta = \begin{bmatrix} \text{var}(\delta_1) & & & & & & \\ 0 & \text{var}(\delta_2) & & & & & \\ 0 & 0 & \text{var}(\delta_3) & & & & \\ 0 & 0 & 0 & \text{var}(\delta_4) & & & \\ 0 & \phi_{52} & 0 & 0 & \text{var}(\delta_5) & & \\ 0 & 0 & 0 & 0 & 0 & \text{var}(\delta_6) \end{bmatrix}$$

I fattori specifici sono incorrelati!

L'Analisi Fattoriale Confermativa: un esempio

$$\Sigma = \Lambda_x \Phi \Lambda'_x + \Theta$$

$$\Sigma = \begin{bmatrix} \text{var}(x_1) & & & & & \\ \text{cov}(x_2, x_1) & \text{var}(x_2) & & & & \\ \text{cov}(x_3, x_1) & \text{cov}(x_3, x_2) & \text{var}(x_3) & & & \\ \text{cov}(x_4, x_1) & \text{cov}(x_4, x_2) & \text{cov}(x_4, x_3) & \text{var}(x_4) & & \\ \text{cov}(x_5, x_1) & \text{cov}(x_5, x_2) & \text{cov}(x_5, x_4) & \text{cov}(x_5, x_4) & \text{var}(x_5) & \\ \text{cov}(x_6, x_1) & \text{cov}(x_6, x_2) & \text{cov}(x_6, x_4) & \text{cov}(x_6, x_4) & \text{cov}(x_6, x_5) & \text{var}(x_6) \end{bmatrix}$$

$$\Sigma(\Omega) = \begin{bmatrix} \lambda_{11}^2 \phi_{11} + \text{var}(\delta_1) & & & & & \\ \lambda_{11} \lambda_{21} \phi_{11} + \lambda_{11} \lambda_{22} \phi_{21} & \lambda_{21}^2 \phi_{11} + 2 \lambda_{21} \lambda_{22} \phi_{21} + \lambda_{21}^2 \phi_{22} + \text{var}(\delta_2) & & & & \\ \lambda_{11} \lambda_{31} \phi_{11} & \lambda_{21} \lambda_{31} \phi_{11} + \lambda_{22} \lambda_{31} \phi_{21} & \lambda_{31}^2 \phi_{11} + \text{var}(\delta_3) & & & \\ \lambda_{11} \lambda_{42} \phi_{21} & \lambda_{21} \lambda_{42} \phi_{21} + \lambda_{22} \lambda_{42} \phi_{22} & \lambda_{31} \lambda_{42} \phi_{21} & \lambda_{42}^2 \phi_{22} + \text{var}(\delta_4) & & \\ \lambda_{11} \lambda_{52} \phi_{21} & \lambda_{21} \lambda_{52} \phi_{21} + \lambda_{22} \lambda_{52} \phi_{22} & \lambda_{31} \lambda_{52} \phi_{21} & \lambda_{42} \lambda_{52} \phi_{21} & \lambda_{52}^2 \phi_{22} + \text{var}(\delta_5) & \\ \lambda_{11} \lambda_{62} \phi_{21} & \lambda_{21} \lambda_{62} \phi_{21} + \lambda_{22} \lambda_{62} \phi_{22} & \lambda_{31} \lambda_{62} \phi_{21} & \lambda_{42} \lambda_{62} \phi_{22} & \lambda_{52} \lambda_{62} \phi_{22} & \lambda_{62}^2 \phi_{22} + \text{var}(\delta_6) \end{bmatrix}$$

L'Analisi Fattoriale Confermativa: un esempio

$$\Sigma = \Lambda_x \Phi \Lambda'_x + \Theta$$

ad esempio:

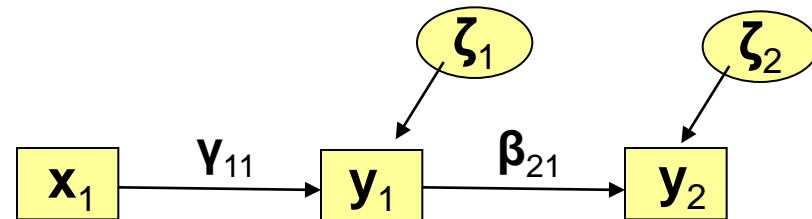
$$\text{var}(x_3) = \lambda_{31}^2 \phi_{11} + \text{var}(\delta_3)$$

La varianza della variabile x_3 è funzione della varianza del fattore comune (ξ_1), del factor loadings che la lega al suo fattore comune (λ_{31}) e della varianza del fattore specifico (δ_3), in altre parole è funzione dei parametri del modello.

Un Modello ad Equazioni Strutturali con solo VM

Un modello di regressione multipla e/o un sistema di regressioni semplici o multiple possono essere visti come un caso particolare dei Modelli ad Equazioni Strutturali, nel quale tutte le variabili (tranne l'errore) sono direttamente osservate

$$y_1 = \gamma_{11} x_1 + \zeta_1$$
$$y_2 = \beta_{21} y_1 + \zeta_2$$

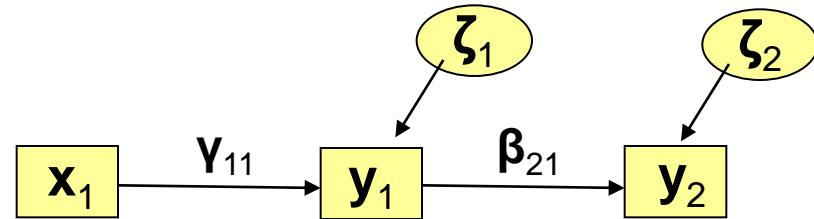


$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0 & 0 \\ \beta_{21} & 0 \end{bmatrix} \quad \mathbf{\Gamma} = \begin{bmatrix} \gamma_{11} \\ 0 \end{bmatrix} \quad \boldsymbol{\zeta} = \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}$$

$$\mathbf{y} = \mathbf{\Gamma} \mathbf{x} + \mathbf{B} \mathbf{y} + \boldsymbol{\zeta}$$

Un Modello ad Equazioni Strutturali con solo VM

$$\mathbf{y} = \Gamma \mathbf{x} + \mathbf{B} \mathbf{y} + \boldsymbol{\zeta}$$



$$\mathbf{y} = (\mathbf{I} - \mathbf{B})^{-1} (\Gamma \mathbf{x} + \boldsymbol{\zeta})$$

Assumendo che:

→ la covarianza tra due errori sia nulla

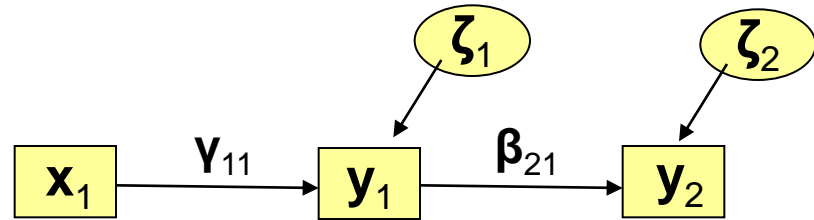
→ le variabili endogene e gli errori non covarino tra loro

e che:

$$\Sigma = \begin{bmatrix} \Sigma_{xx} & \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

Un Modello ad Equazioni Strutturali con solo VM

$$\mathbf{y} = (\mathbf{I} - \mathbf{B})^{-1} (\mathbf{\Gamma} \mathbf{x} + \boldsymbol{\zeta})$$



$$\Sigma_{xx}(\Omega) = \Phi = E(\mathbf{x}\mathbf{x}')$$

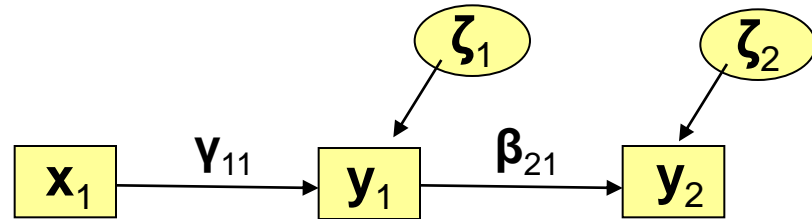
e

$$\Sigma_{yx}(\Omega) = E(\mathbf{x}\mathbf{y}') = E\left[\mathbf{x}\left((\mathbf{I} - \mathbf{B})^{-1}(\mathbf{\Gamma}\mathbf{x} + \boldsymbol{\zeta})\right)'\right]$$
$$= \Phi\mathbf{\Gamma}'(\mathbf{I} - \mathbf{B})^{-1'}$$

Un Modello ad Equazioni Strutturali con solo VM

$$\mathbf{y} = (\mathbf{I} - \mathbf{B})^{-1} (\mathbf{\Gamma}\mathbf{x} + \boldsymbol{\zeta})$$

con



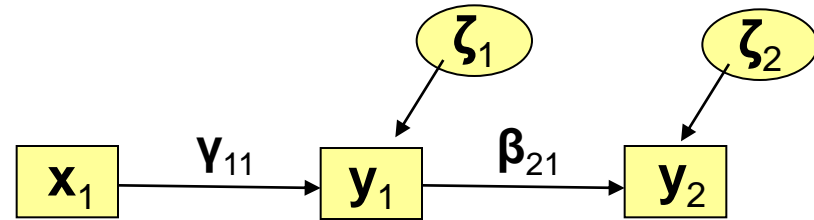
$$\Phi = E(\mathbf{x}\mathbf{x}') \quad e \quad \Psi = E(\boldsymbol{\zeta}\boldsymbol{\zeta}')$$

Si avrà:

$$\begin{aligned} \Sigma_{yy}(\Omega) &= E(\mathbf{y}\mathbf{y}') = E\left[(\mathbf{I} - \mathbf{B})^{-1} (\mathbf{\Gamma}\mathbf{x} + \boldsymbol{\zeta}) \left((\mathbf{I} - \mathbf{B})^{-1} (\mathbf{\Gamma}\mathbf{x} + \boldsymbol{\zeta}) \right)' \right] \\ &= E\left[(\mathbf{I} - \mathbf{B})^{-1} (\mathbf{\Gamma}\mathbf{x} + \boldsymbol{\zeta}) (\mathbf{x}\boldsymbol{\Gamma}' + \boldsymbol{\zeta}') (\mathbf{I} - \mathbf{B})^{-1'} \right] \\ &= (\mathbf{I} - \mathbf{B})^{-1} \left[E(\boldsymbol{\Gamma}\mathbf{x}\mathbf{x}\boldsymbol{\Gamma}') + E(\boldsymbol{\Gamma}\mathbf{x}\boldsymbol{\zeta}') + E(\boldsymbol{\zeta}\mathbf{x}\boldsymbol{\Gamma}') + E(\boldsymbol{\zeta}\boldsymbol{\zeta}') \right] (\mathbf{I} - \mathbf{B})^{-1'} \\ &= (\mathbf{I} - \mathbf{B})^{-1} (\boldsymbol{\Gamma}\Phi\boldsymbol{\Gamma}' + \Psi) (\mathbf{I} - \mathbf{B})^{-1'} \end{aligned}$$

Un Modello ad Equazioni Strutturali con solo VM

$$\mathbf{y} = (\mathbf{I} - \mathbf{B})^{-1} (\mathbf{\Gamma} \mathbf{x} + \boldsymbol{\zeta})$$



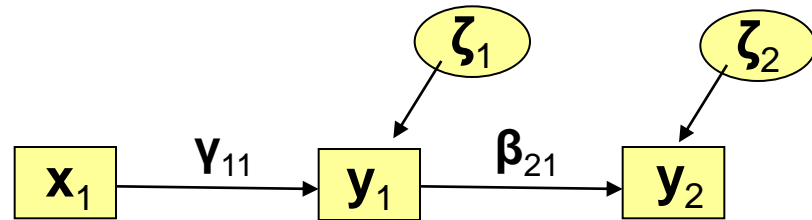
Sulla base dei risultati precedenti:

$$\Sigma = \begin{bmatrix} \Sigma_{xx} & \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

$$\Sigma(\Omega) = \begin{bmatrix} \Phi & \\ (\mathbf{I} - \mathbf{B})^{-1} \Phi \mathbf{\Gamma} & (\mathbf{I} - \mathbf{B})^{-1} (\mathbf{\Gamma} \Phi \mathbf{\Gamma}' + \Psi) (\mathbf{I} - \mathbf{B})^{-1'} \end{bmatrix}$$

Un Modello ad Equazioni Strutturali con solo VM

$$\mathbf{y} = (\mathbf{I} - \mathbf{B})^{-1} (\mathbf{\Gamma} \mathbf{x} + \boldsymbol{\zeta})$$



Ricordando che:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0 & 0 \\ \beta_{21} & 0 \end{bmatrix} \quad \mathbf{\Gamma} = \begin{bmatrix} \gamma_{11} \\ 0 \end{bmatrix} \quad \boldsymbol{\zeta} = \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}$$

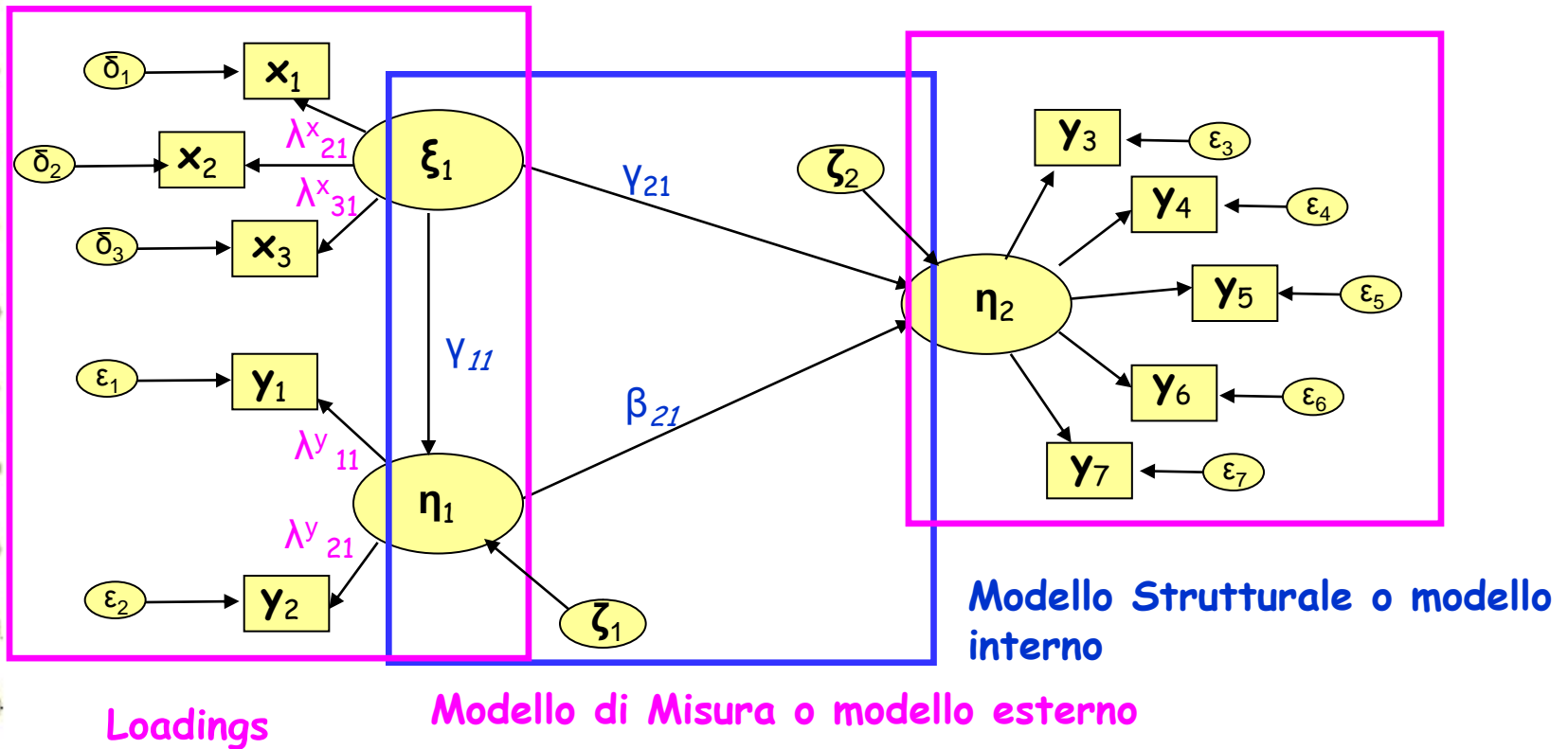
$$\Sigma(\Omega) = \begin{bmatrix} \phi_{11} & & \\ \gamma_{11}\phi_{11} & \gamma_{11}^2\phi_{11} + \psi_{11} & \\ \beta_{21}\gamma_{11}\phi_{11} & \beta_{21}(\gamma_{11}^2\phi_{11} + \psi_{11}) & \beta_{21}^2(\gamma_{11}^2\phi_{11} + \psi_{11}) + \psi_{22} \end{bmatrix}$$

Modelli ad Equazioni Strutturali - corso 2009

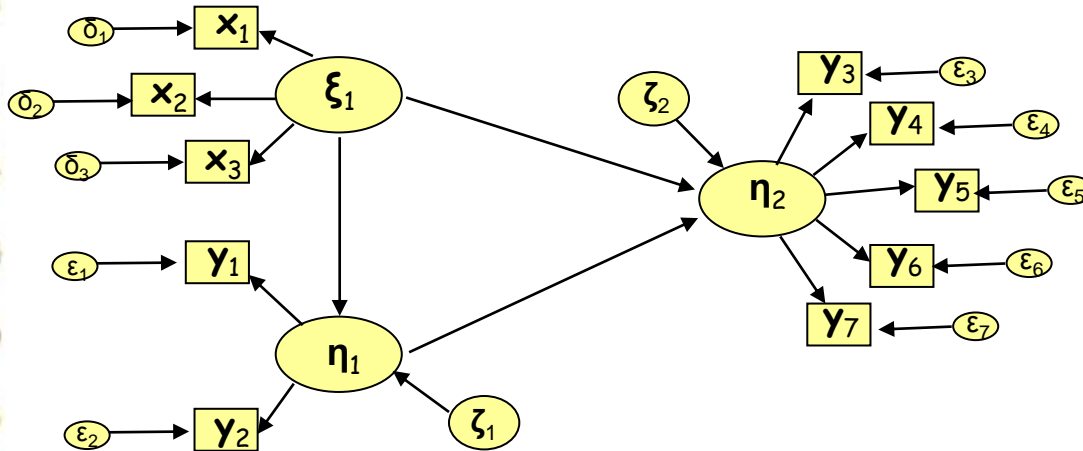
Lezione 2

Modelli ad Equazioni Strutturali: le nozioni di base

Path
Coefficients



Modelli ad Equazioni Strutturali: le nozioni di base



- * **P Variabili Manifeste (VM) Esogene** $\rightarrow x_p$ generica VM esogena
e **P Errori di misura** $\rightarrow \delta_p$ generico errore associato alla VM esogena
- * **Q Variabili Manifeste (VM) Endogene** $\rightarrow y_q$ generica VM endogena
e **Q Errori di misura** $\rightarrow \varepsilon_q$ generico errore associato alla VM endogena
- * **M Variabili Latenti (VL) Esogene** $\rightarrow \xi_m$ generica VL esogena
- * **J Variabili Latenti (VL) Endogene** $\rightarrow \eta_j$ generica VL endogena
e **J Errori strutturali** $\rightarrow \zeta_j$ generico errore strutturale

Modelli ad Equazioni Strutturali: un esempio

→ $P=3 \quad Q=7 \quad M=1 \quad J=2$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_3 \end{bmatrix}$$

$$\boldsymbol{\delta} = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_3 \end{bmatrix}$$

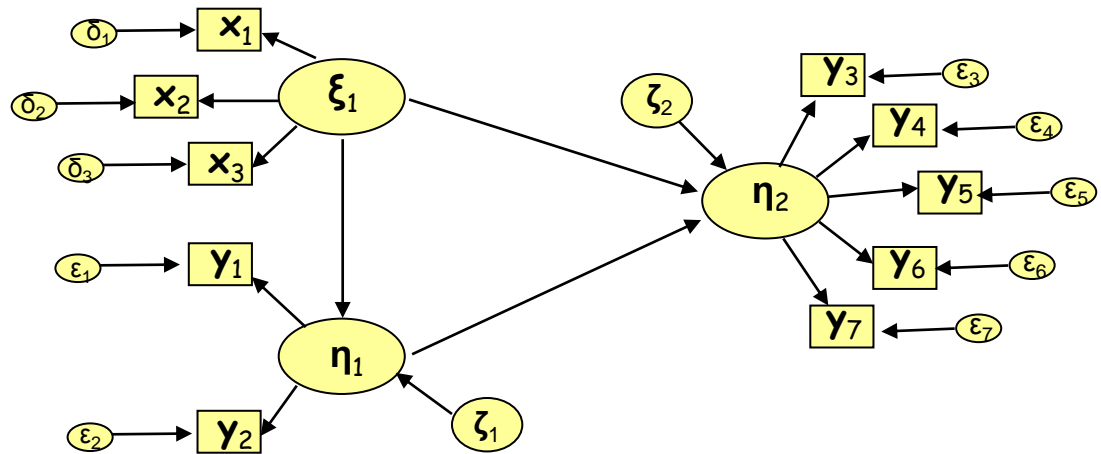
$$\boldsymbol{\xi} = [\xi_1]$$

$$\boldsymbol{\eta} = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$$

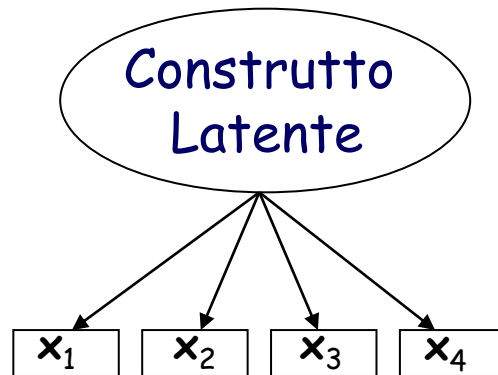
$$\boldsymbol{\zeta} = \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_7 \end{bmatrix}$$

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_7 \end{bmatrix}$$



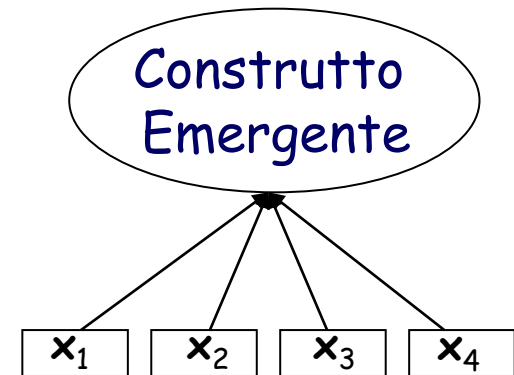
Modelli ad Equazioni Strutturali: il modello di misura



Indicatori Riflessivi

es. Intelligenza

- Le VL **generano le corrispondenti VM** (la VL sottostante ogni blocco è unica → blocco unidimensionale)
- Le VL sono **antecedenti rispetto alle VM**
- Gli indicatori riflessivi di uno stesso blocco devono essere fortemente legati tra loro, in altre parole devono **covariare**
- **La consistenza Interna al blocco** deve essere verificata (ad esempio usando l'alpha di Cronbach)



Indicatori Formativi

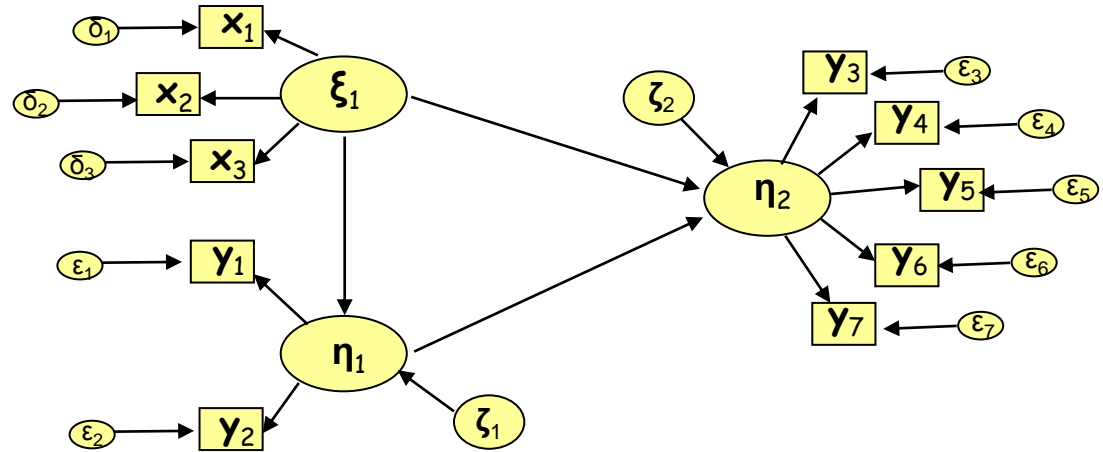
es. Status Sociale

- La VL è una **combinazione lineare delle corrispondenti VM** (ogni VL è quindi un costrutto multidimensionale)
- Non è detto che gli indicatori formativi di uno stesso blocco debbano covariare tra loro.
- La consistenza Interna non necessita di essere verificata.

Modelli ad Equazioni Strutturali: il modello di misura

Per le VM esogene

$$\begin{aligned} x_1 &= \lambda_{11}^x \xi_1 + \delta_1 \\ x_2 &= \lambda_{21}^x \xi_1 + \delta_2 \\ x_3 &= \lambda_{31}^x \xi_1 + \delta_3 \end{aligned}$$



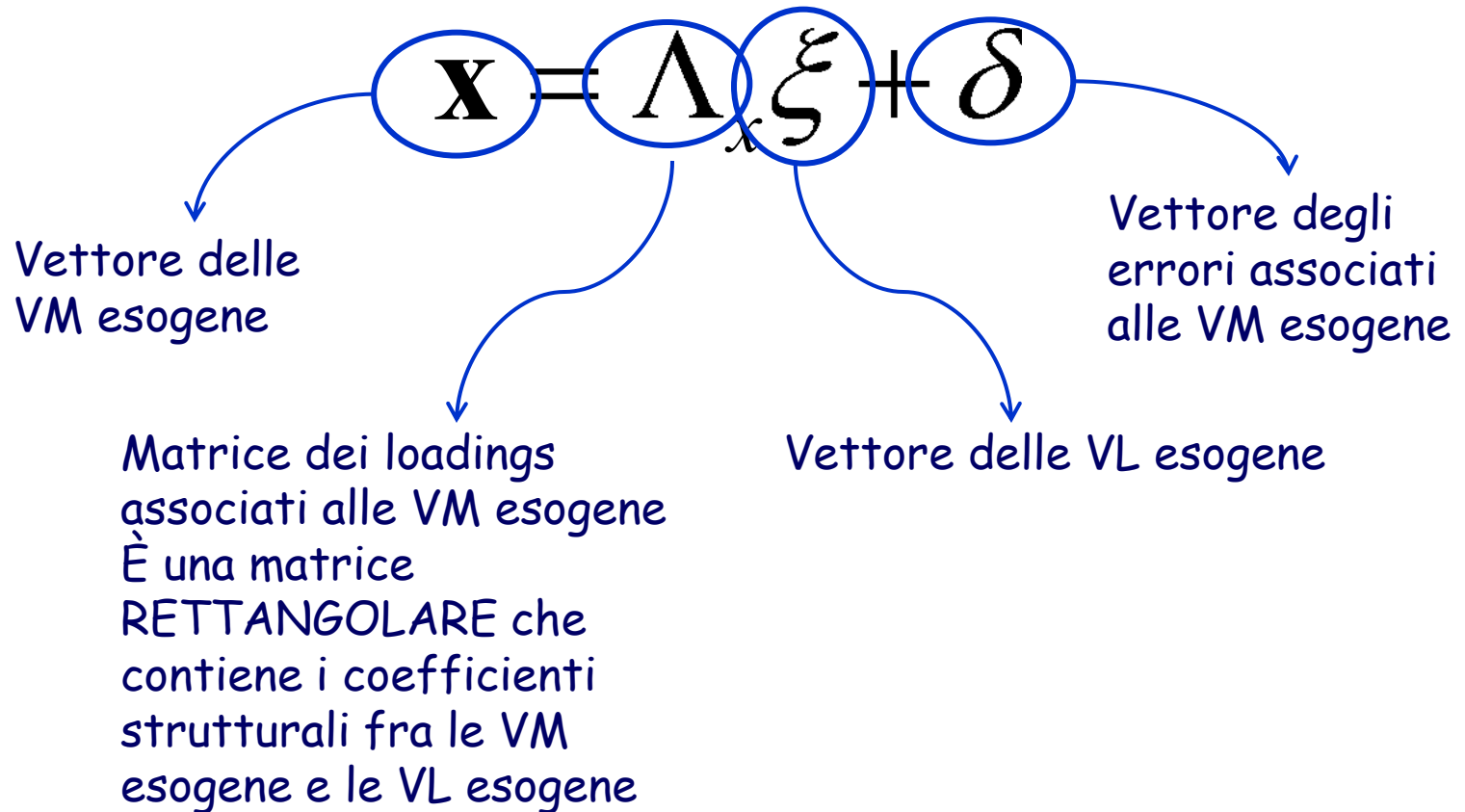
$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \Lambda_x = \begin{bmatrix} \lambda_{11}^x \\ \lambda_{21}^x \\ \lambda_{31}^x \end{bmatrix} \quad \xi = \begin{bmatrix} \xi_1 \end{bmatrix} \quad \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix}$$

$(P \times 1)$ $(P \times M)$ $(M \times 1)$ $(P \times 1)$

$$\mathbf{x} = \Lambda_x \xi + \delta$$

Modelli ad Equazioni Strutturali: il modello di misura

Per le VM esogene



Modelli ad Equazioni Strutturali: il modello di misura

Per le VM endogene

$$y_1 = \lambda_{11}^y \eta_1 + \varepsilon_1$$

$$y_2 = \lambda_{21}^y \eta_1 + \varepsilon_2$$

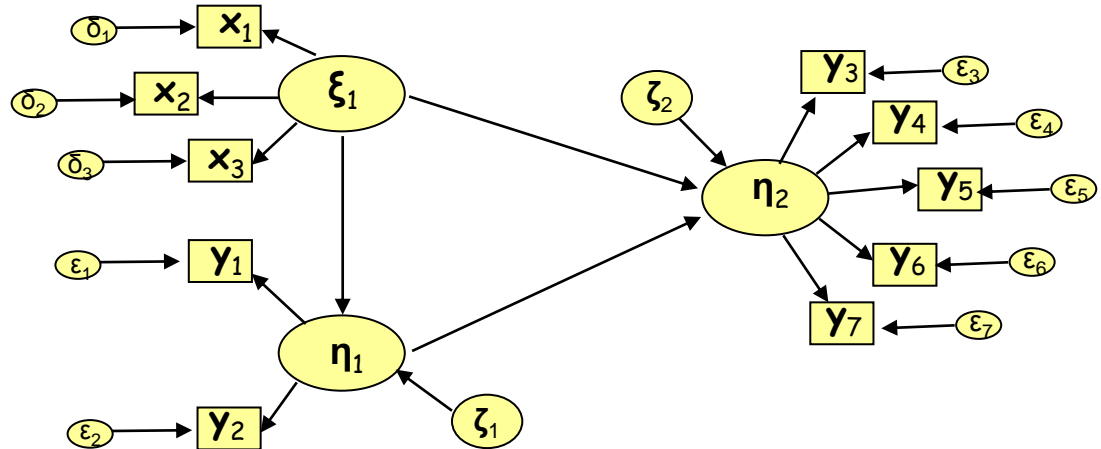
$$y_3 = \lambda_{32}^y \eta_2 + \varepsilon_3$$

$$y_4 = \lambda_{42}^y \eta_2 + \varepsilon_4$$

$$y_5 = \lambda_{52}^y \eta_2 + \varepsilon_5$$

$$y_6 = \lambda_{62}^y \eta_2 + \varepsilon_6$$

$$y_7 = \lambda_{72}^y \eta_2 + \varepsilon_7$$

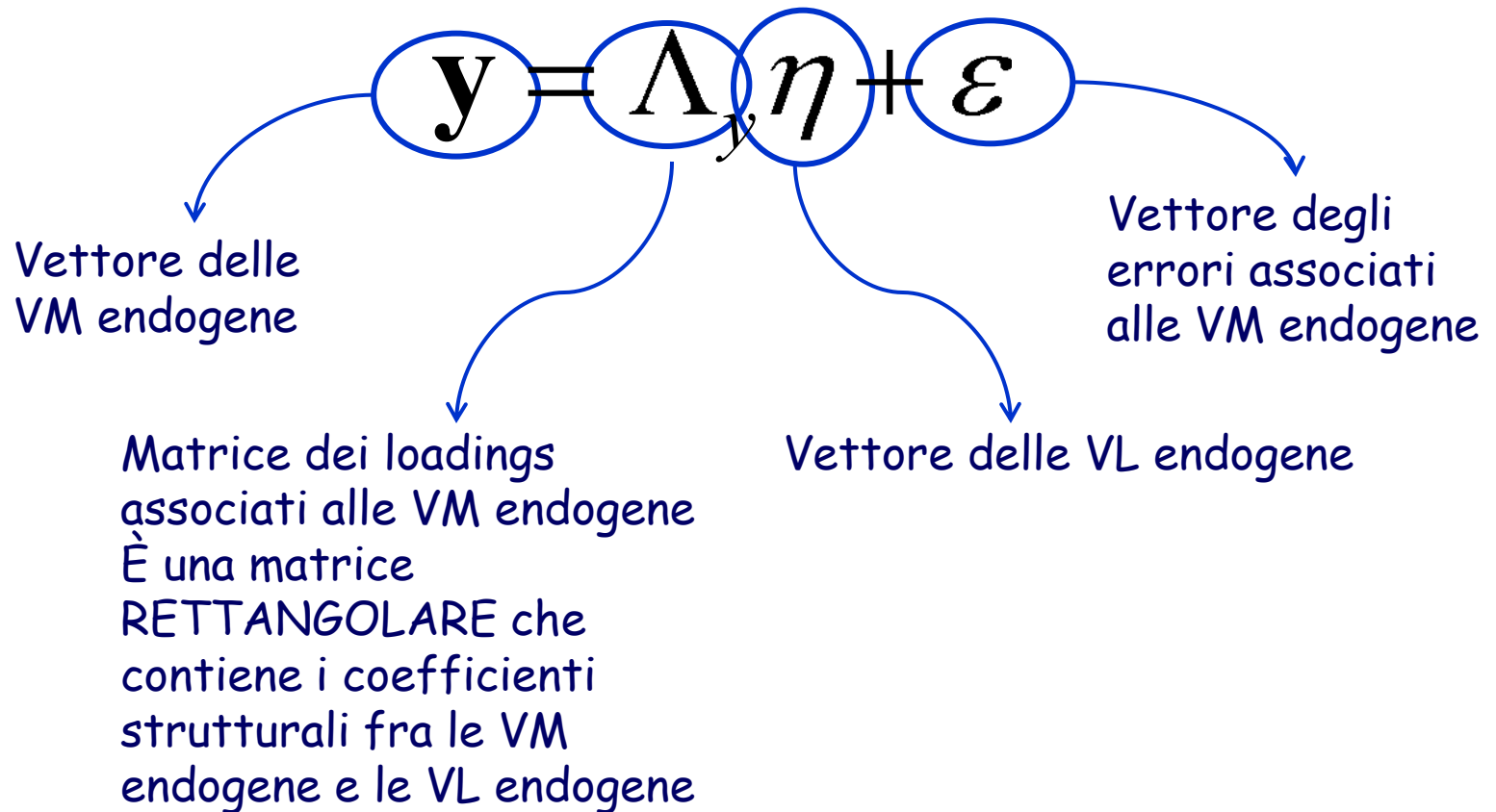


$$\begin{matrix} \mathbf{y} = \\ (Q \times 1) \end{matrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{bmatrix} \quad \Lambda_y = \begin{matrix} (Q \times J) \end{matrix} \begin{bmatrix} \lambda_{11}^y & 0 \\ \lambda_{21}^y & 0 \\ 0 & \lambda_{32}^y \\ 0 & \lambda_{42}^y \\ 0 & \lambda_{52}^y \\ 0 & \lambda_{62}^y \\ 0 & \lambda_{72}^y \end{bmatrix} \quad \eta = \begin{matrix} (J \times 1) \end{matrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \quad \varepsilon = \begin{matrix} (Q \times 1) \end{matrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \end{bmatrix}$$

$$\mathbf{y} = \Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

Modelli ad Equazioni Strutturali: il modello di misura

Per le VM endogene

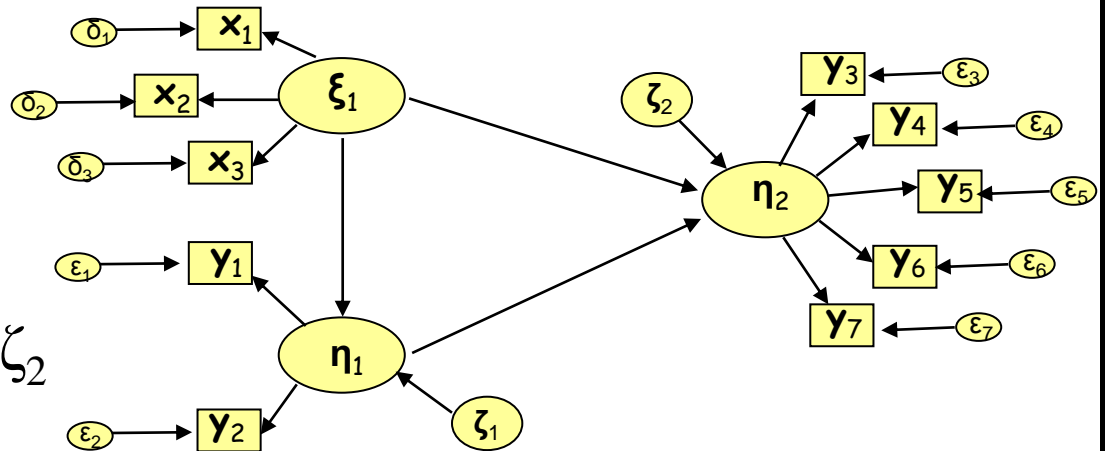


Modelli ad Equazioni Strutturali: il modello strutturale

Per ogni VL endogena nel modello

$$\eta_1 = \gamma_{11} \xi_1 + \zeta_1$$

$$\eta_2 = \gamma_{21} \xi_1 + \beta_{21} \eta_1 + \zeta_2$$



$$\eta = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \quad (J \times 1)$$

$$\xi = [\xi_1] \quad (M \times 1)$$

$$\Gamma = \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \end{bmatrix} \quad (J \times M)$$

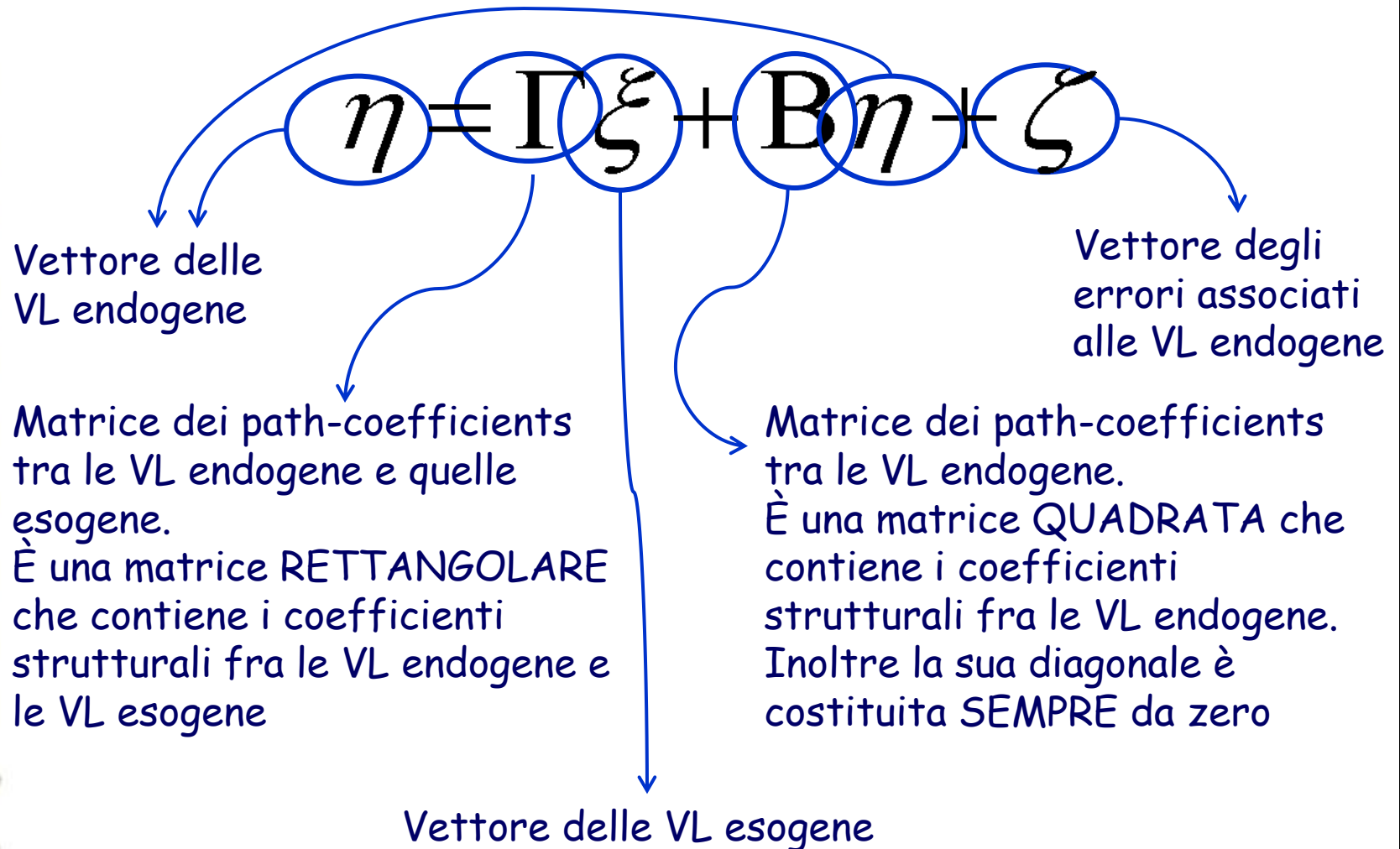
$$B = \begin{bmatrix} 0 & 0 \\ \beta_{21} & 0 \end{bmatrix} \quad (J \times J)$$

$$\zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} \quad (J \times 1)$$

$$\eta = \Gamma \xi + B \eta + \zeta$$

Modelli ad Equazioni Strutturali: il modello strutturale

Per le VL endogene



Modelli ad Equazioni Strutturali: il modello strutturale

Per le VL endogene

$$\eta = \Gamma \xi + \mathbf{B} \eta + \zeta$$

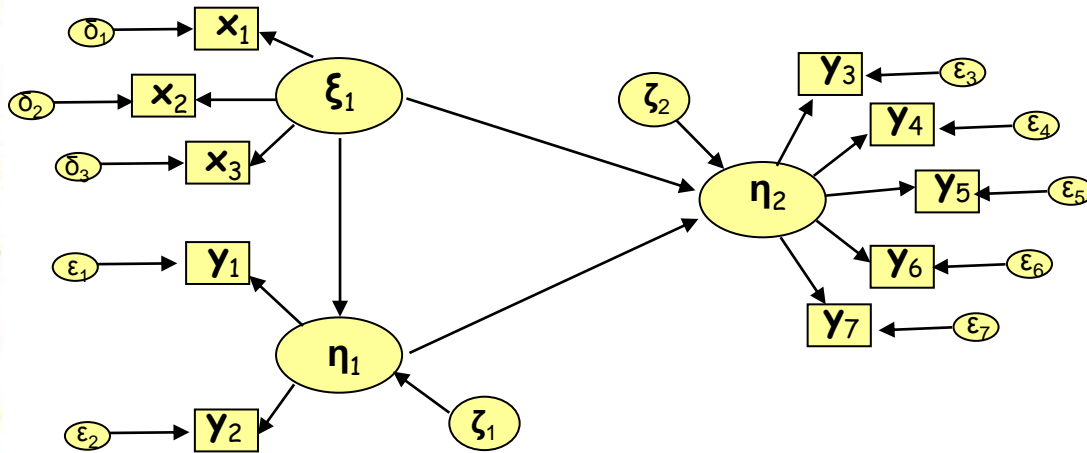
$$(\mathbf{I} - \mathbf{B}) \eta = \Gamma \xi + \zeta$$

$$\eta = (\mathbf{I} - \mathbf{B})^{-1} (\Gamma \xi + \zeta)$$

posto $\tilde{\mathbf{B}} = (\mathbf{I} - \mathbf{B})$

$$\eta = \tilde{\mathbf{B}}^{-1} (\Gamma \xi + \zeta)$$

Modelli ad Equazioni Strutturali: le tre equazioni



$$\mathbf{x} = \Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta}$$

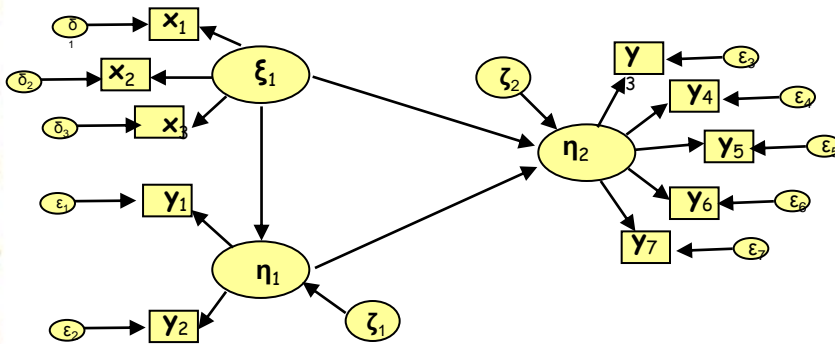
Modello di Misura

$$\mathbf{y} = \Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

Modello Strutturale

$$\boldsymbol{\eta} = \boldsymbol{\Gamma} \boldsymbol{\xi} + \mathbf{B} \boldsymbol{\eta} + \boldsymbol{\zeta} \leftrightarrow \boldsymbol{\eta} = (\mathbf{I} - \mathbf{B})^{-1} (\boldsymbol{\Gamma} \boldsymbol{\xi} + \boldsymbol{\zeta})$$

Modelli ad Equazioni Strutturali: le assunzioni



$$\mathbf{x} = \Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta}$$

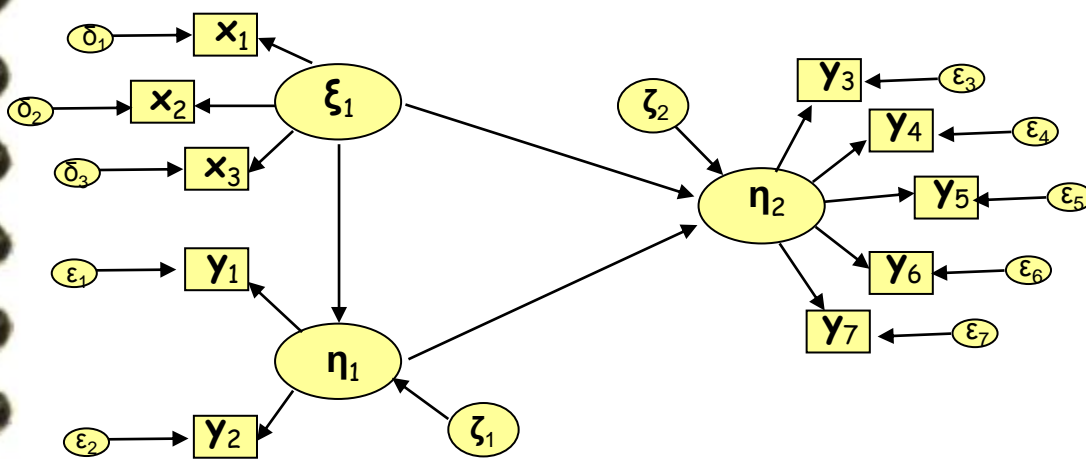
$$\mathbf{y} = \Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\eta} = (\mathbf{I} - \mathbf{B})^{-1} (\boldsymbol{\Gamma} \boldsymbol{\xi} + \boldsymbol{\zeta})$$

Si assume che:

- i) le VM, le VL e gli errori (sia del modello strutturale che del modello di misura) sono centrate
- ii) la covarianza tra due errori (sia del modello strutturale che del modello di misura) sia nulla
- iii) gli errori di misura (esogeni ed endogeni) e le VL (esogene ed endogene) non covarino tra loro
- iv) gli errori strutturali e le VL esogene non covarino tra loro
- v) il modello strutturale deve essere NON ridondante

Modelli ad Equazioni Strutturali: le assunzioni



$$\mathbf{x} = \Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta}$$

$$\mathbf{y} = \Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\eta} = (\mathbf{I} - \mathbf{B})^{-1} (\boldsymbol{\Gamma} \boldsymbol{\xi} + \boldsymbol{\zeta})$$

In altre parole si assume che:

i) $E(\mathbf{x}) = 0, E(\mathbf{y}) = 0, E(\boldsymbol{\xi}) = 0, E(\boldsymbol{\eta}) = 0, E(\boldsymbol{\delta}) = 0, E(\boldsymbol{\varepsilon}) = 0, E(\boldsymbol{\zeta}) = 0$

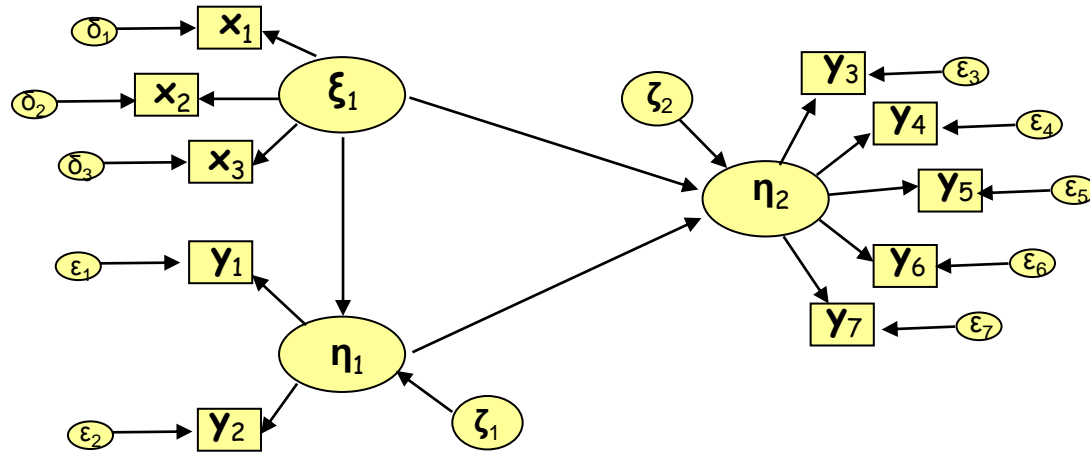
ii) $E(\boldsymbol{\delta} \boldsymbol{\varepsilon}') = 0, E(\boldsymbol{\zeta} \boldsymbol{\delta}') = 0, E(\boldsymbol{\zeta} \boldsymbol{\varepsilon}') = 0$

iii) $E(\boldsymbol{\delta} \boldsymbol{\xi}') = 0, E(\boldsymbol{\delta} \boldsymbol{\eta}') = 0, E(\boldsymbol{\varepsilon} \boldsymbol{\xi}') = 0, E(\boldsymbol{\varepsilon} \boldsymbol{\eta}') = 0$

iv) $E(\boldsymbol{\zeta} \boldsymbol{\xi}') = 0$

v) $(\mathbf{I} - \mathbf{B})$ sia non singolare

Modelli ad Equazioni Strutturali: la matrice di var/cov



Data la matrice di var/cov tra le VM :

Matrice var/cov tra le VM esogene $(P \times P)$

Matrice di di intercovarianza tra le VM endogene e le VM esogene $(Q \times P)$

$$\Sigma = \begin{bmatrix} \Sigma_{xx} \\ \Sigma_{yx} \end{bmatrix} \begin{bmatrix} \Sigma_{yy} \end{bmatrix}$$

$((P+Q) \times (P+Q))$

Matrice var/cov tra le VM endogene $(Q \times Q)$

Modelli ad Equazioni Strutturali: la matrice Σ_{XX}

Ricordando che il modello di misura per le variabili esogene si scrive:

$$\mathbf{x} = \Lambda_x \xi + \delta$$

La matrice di var/cov tra le VM esogene può essere riscritta in termini di parametri del modello, come:

$$\begin{aligned}\Sigma_{XX}(\Omega) &= E(\mathbf{xx}') = E\left[(\Lambda_x \xi + \delta)(\Lambda_x \xi + \delta)'\right] = \\ &= E\left[(\Lambda_x \xi + \delta)(\xi' \Lambda_x' + \delta')\right] = \\ &= \Lambda_x E(\xi \xi') \Lambda_x' + \Lambda_x E(\xi \delta') + E(\delta \xi') \Lambda_x' + E(\delta \delta')\end{aligned}$$

posto $E(\xi \xi') = \Phi$
 $E(\delta \delta') = \Theta_\delta$



$$\Sigma_{XX}(\Omega) = \Lambda_x \Phi \Lambda_x' + \Theta_\delta$$

Modelli ad Equazioni Strutturali: la matrice Σ_{XX}

$$\Sigma_{XX}(\Omega) = \Lambda_x \Phi \Lambda_x' + \Theta_\delta$$

È la matrice var/cov tra le VM esogene espressa in termini di parametri del modello
($P \times P$)

Modelli ad Equazioni Strutturali: la matrice Σ_{YY}

Ricordando che il modello di misura per le variabili endogene si scrive:

$$\mathbf{y} = \Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

La matrice di var/cov tra le VM endogene può essere riscritta in termini di parametri del modello, come:

$$\begin{aligned}\Sigma_{YY}(\boldsymbol{\Omega}) &= E(\mathbf{y}\mathbf{y}') = E\left[(\Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon})(\Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon})'\right] = \\ &= E\left[(\Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon})(\boldsymbol{\eta}'\Lambda_y' + \boldsymbol{\varepsilon}')\right] = \\ &= E\left[\Lambda_y \boldsymbol{\eta}\boldsymbol{\eta}'\Lambda_y' + \boldsymbol{\varepsilon}\boldsymbol{\eta}'\Lambda_y' + \Lambda_y \boldsymbol{\eta}\boldsymbol{\varepsilon}' + \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right] = \\ &= \Lambda_y E(\boldsymbol{\eta}\boldsymbol{\eta}')\Lambda_y' + \Lambda_y E(\boldsymbol{\varepsilon}\boldsymbol{\eta}') + E(\boldsymbol{\eta}\boldsymbol{\varepsilon}')\Lambda_y' + E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\end{aligned}$$

posto $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \Theta_\varepsilon \quad \rightarrow \quad \Sigma_{YY}(\boldsymbol{\Omega}) = \Lambda_y E(\boldsymbol{\eta}\boldsymbol{\eta}')\Lambda_y' + \Theta_\varepsilon$

Modelli ad Equazioni Strutturali: la matrice Σ_{YY}

Ricordando che il modello strutturale esprime le VL endogene come:

$$\eta = (\mathbf{I} - \mathbf{B})^{-1} (\Gamma \xi + \zeta)$$

La matrice di var/cov tra le VL endogene può essere riscritta come:

$$\Sigma_{\eta\eta} = E(\eta\eta') = E\left(\left[(\mathbf{I} - \mathbf{B})^{-1} (\Gamma \xi + \zeta) \right] \left[(\mathbf{I} - \mathbf{B})^{-1} (\Gamma \xi + \zeta) \right]' \right) =$$

$$= E\left(\left[(\mathbf{I} - \mathbf{B})^{-1} (\Gamma \xi + \zeta) \right] (\xi \Gamma' + \zeta \zeta') (\mathbf{I} - \mathbf{B})^{-1'} \right) =$$

$$= (\mathbf{I} - \mathbf{B})^{-1} E(\Gamma \xi \xi \Gamma' + \zeta \xi \Gamma' + \Gamma \xi \zeta' + \zeta \zeta') (\mathbf{I} - \mathbf{B})^{-1'} =$$

$$= (\mathbf{I} - \mathbf{B})^{-1} \left[\Gamma E(\xi \xi') \Gamma' + E(\zeta \xi') \Gamma' + \Gamma E(\xi \zeta') + E(\zeta \zeta') \right] (\mathbf{I} - \mathbf{B})^{-1'}$$

posto $E(\zeta \zeta') = \Psi \rightarrow \Sigma_{\eta\eta}(\Omega) = (\mathbf{I} - \mathbf{B})^{-1} (\Gamma \Phi \Gamma' + \Psi) (\mathbf{I} - \mathbf{B})^{-1'}$

Modelli ad Equazioni Strutturali: la matrice Σ_{YY}

Sostituendo $\Sigma_{\eta\eta}(\Omega) = (\mathbf{I} - \mathbf{B})^{-1} (\Gamma\Phi\Gamma' + \Psi)(\mathbf{I} - \mathbf{B})^{-1'}$

in $\Sigma_{YY}(\Omega) = \Lambda_y E(\eta\eta') \Lambda'_y + \Theta_\varepsilon$



$$\Sigma_{YY}(\Omega) = \Lambda_y \left[(\mathbf{I} - \mathbf{B})^{-1} (\Gamma\Phi\Gamma' + \Psi)(\mathbf{I} - \mathbf{B})^{-1'} \right] \Lambda'_y + \Theta_\varepsilon$$

È la matrice var/cov tra le VM endogene espressa in termini di parametri del modello
($Q \times Q$)

Modelli ad Equazioni Strutturali: la matrice Σ_{XY}

Ricordando che il modello di misura si scrive:

$$\mathbf{x} = \Lambda_x \xi + \delta \quad e \quad \mathbf{y} = \Lambda_y \eta + \varepsilon$$

La matrice di intercovarianza tra le VM endogene e quelle esogene può essere riscritta in termini di parametri del modello, come:

$$\Sigma_{XY}(\Omega) = E(\mathbf{XY}') = E\left[(\Lambda_x \xi + \delta)(\Lambda_y \eta + \varepsilon)'\right] =$$

$$= E\left[(\Lambda_x \xi + \delta)(\eta' \Lambda_y' + \varepsilon')\right] =$$

$$= \Lambda_x E(\xi \eta') \Lambda_y' + \Lambda_x E(\xi \varepsilon') + E(\delta \eta') \Lambda_y' + E(\delta \varepsilon')$$

Ricordando che gli errori sono incorrelati tra loro e con le VL esogene

$$\rightarrow \Sigma_{XY}(\Omega) = \Lambda_x E(\xi \eta') \Lambda_y'$$

Modelli ad Equazioni Strutturali: la matrice Σ_{XY}

Ricordando che il modello strutturale esprime le VL endogene come:

$$\eta = (\mathbf{I} - \mathbf{B})^{-1} (\Gamma \xi + \zeta)$$

E sostituendo η nell'equazione precedente, si ha:

$$\Sigma_{XY}(\Omega) = \Lambda_x \mathbf{E}(\xi \eta') \Lambda_y' = \Lambda_x \mathbf{E} \left(\xi \left((\mathbf{I} - \mathbf{B})^{-1} (\Gamma \xi + \zeta) \right)' \right) \Lambda_y' =$$

$$= \Lambda_x \mathbf{E} \left(\xi \left(\xi' \Gamma' (\mathbf{I} - \mathbf{B})^{-1'} + \zeta' (\mathbf{I} - \mathbf{B})^{-1'} \right) \right) \Lambda_y' =$$

$$= \Lambda_x \left[\mathbf{E}(\xi \xi') \Gamma' (\mathbf{I} - \mathbf{B})^{-1'} + \mathbf{E}(\xi \zeta') (\mathbf{I} - \mathbf{B})^{-1'} \right] \Lambda_y'$$

$$\longrightarrow \Sigma_{XY}(\Omega) = \Lambda_x \Phi \Gamma' (\mathbf{I} - \mathbf{B})^{-1'} \Lambda_y'$$

Modelli ad Equazioni Strutturali: la matrice Σ_{XY}

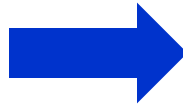
$$\Sigma_{XY}(\Omega) = \Lambda_x \Phi \Gamma' (\mathbf{I} - \mathbf{B})^{-1'} \Lambda_y'$$

È la matrice di di intercovarianza tra le VM endogene e le VM esogene espressa in termini di parametri del modello
($P \times Q$)

$$\Sigma_{YX} = \Sigma_{XY}' = \Lambda_y (\mathbf{I} - \mathbf{B})^{-1} \Gamma \Phi' \Lambda_x'$$

Modelli ad Equazioni Strutturali: la matrice di var/cov

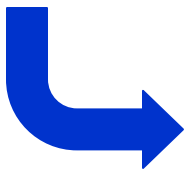
$$\Sigma = \begin{bmatrix} \Sigma_{xx} & \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$



Matrice di var/cov
nella popolazione

Può essere riscritta in termini di parametri del modello:

$$\Sigma(\Omega) = \begin{bmatrix} \Lambda_x \Phi \Lambda_x' + \Theta_\delta \\ \Lambda_y (\mathbf{I} - \mathbf{B})^{-1} \Gamma \Phi' \Lambda_x' \\ \Lambda_y \left[(\mathbf{I} - \mathbf{B})^{-1} (\Gamma \Phi \Gamma' + \Psi) (\mathbf{I} - \mathbf{B})^{-1'} \right] \Lambda_y' + \Theta_\varepsilon \end{bmatrix}$$



Matrice di var/cov "attesa" o
"implicita" dal modello

Modelli ad Equazioni Strutturali

Per concludere, in ogni Modello ad Equazioni sono presenti

* 4 tipi di variabili

- le VM esogene (X)
- le VM endogene (Y)
- le VL esogene (ξ)
- le VL endogene (η)

* 4 matrici dei coefficienti

- la matrice dei coef. del modello di misura per le VM esogene (Λ_x)
- la matrice dei coef. del modello di misura per le VM endogene (Λ_y)
- la matrice dei coef. Strutturali tra le VL esogene e le VL endogene (Γ)
- la matrice dei coef. Strutturali tra le VL esogene e le VL endogene (B)

* 4 matrici di covarianza

- la matrice di var/cov tra le VL esogene (Φ)
- la matrice di var/cov tra gli errori del modello strutturale (Ψ)
- la matrice di var/cov tra gli errori associati alle VM esogene (Θ_δ)
- la matrice di var/cov tra gli errori associati alle VM endogene (Θ_ϵ)

Metodi di stima dei SEM

Metodi
Covariance-based



Lo scopo è quello di riprodurre la matrice di var/cov delle variabili manifeste osservata attraverso i parametri del modello:

- la matrice di var/cov ottenuta dal modello è funzione dei parametri del modello!!
- è soprattutto un approccio di tipo **confermativo** che viene utilizzato per validare un prestabilito modello (theory building)

SEM

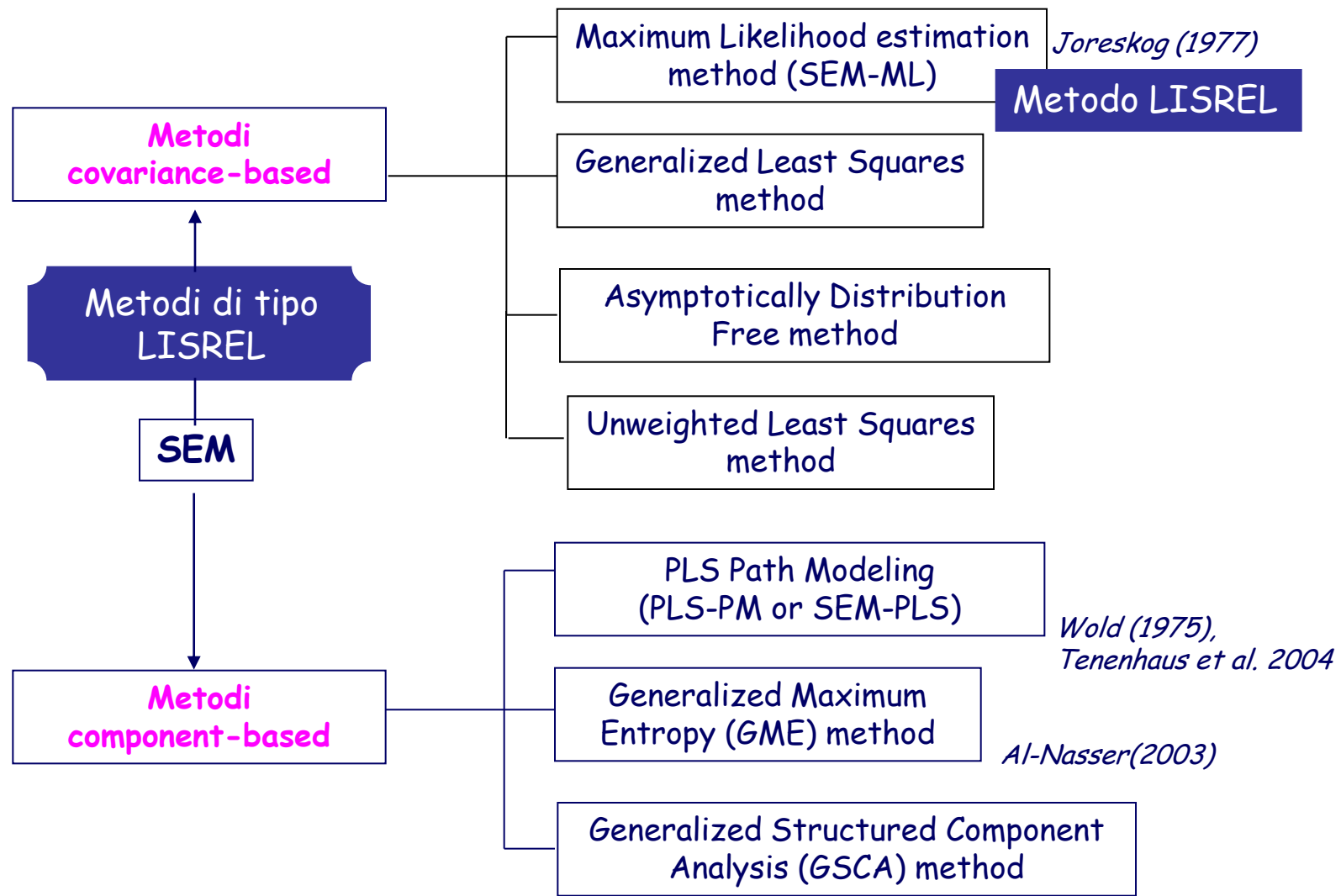
Metodi
Component-based



Lo scopo è quello di fornire una stima delle variabili latenti in modo che esse siano il più rappresentative possibile del proprio blocco di variabili manifeste e contemporaneamente in grado di spiegare al meglio le relazioni di causalità del modello strutturale.

- la stima delle variabili latenti gioca un ruolo centrale nei diversi
- è soprattutto un approccio di tipo **esplorativo** (operational model strategy)

Metodi di stima dei SEM



Joreskog (1977)
Metodo LISREL

Wold (1975), Tenenhaus et al. 2004

Al-Nasser(2003)

Hwang & Takane (2004)

Metodi di stima Covariance-based

$$\Sigma = \begin{bmatrix} \Sigma_{xx} & \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

Matrice di var/cov
nella popolazione

$$S = \begin{bmatrix} S_{xx} & \\ S_{yx} & S_{yy} \end{bmatrix}$$

Matrice di var/cov
osservata

$$C \approx S \approx \Sigma$$

$$C = \Sigma(\Omega) = \begin{bmatrix} \Lambda_x \Phi \Lambda'_x + \Theta_\delta & \\ \Lambda_y (\mathbf{I} - \mathbf{B})^{-1} \Gamma \Phi \Lambda'_x & \Lambda_y \left[(\mathbf{I} - \mathbf{B})^{-1} (\Gamma \Phi \Gamma' + \Psi) (\mathbf{I} - \mathbf{B})^{-1'} \right] \Lambda'_y + \Theta_\varepsilon \end{bmatrix}$$

Matrice di var/cov "attesa" o "implicita" dal modello

Identificabilità del modello

Un modello è **IDENTIFICATO** se i suoi parametri sono univocamente determinati.

Un modello ad equazioni strutturali è **IDENTIFICATO** se la matrice di var/cov è unicamente decomposta in funzione dei parametri del modello.

Condizione NECESSARIA (ma non sufficiente):
gradi di libertà (DF) maggiori o diversi da zero $DF \geq 0$

$$\begin{aligned} DF &= \# \text{ equazioni} - \# \text{ incognite} \\ &= \# \text{ var/cov} - \# \text{ parametri da stimare} \end{aligned}$$

$$DF = \left[\frac{1}{2} (P + Q)(P + Q + 1) - t \right] \geq 0$$

→ Parametri da stimare

Identificabilità del modello

Identificazione Perfetta:

Un modello statistico è perfettamente identificato se l'informazione disponibile implica che ci sia un solo valore ottimo per ogni parametro nel modello il cui valore non è fissato a priori.

→ ZERO gradi di libertà

Un modello perfettamente identificato porta ad un fit "triviale" rendendo il test per la bontà di adattamento non interessante.

Sovraidentificazione:

Un modello si dice sovraidentificato se ci sono più equazioni che incognite. Un modello sovraidentificato non ha un fit perfetto e per questo è più interessante di un modello perfettamente identificato.

→ gradi di libertà > 0

Identificabilità del modello

"Rules of Thumb" :

- **Condizione di ordine**: il numero di variabili escluse in un'equazione deve essere maggiore o uguale al numero di equazioni meno uno (condizione necessaria)
- **Condizione di rango**: vincoli sul rango della matrice dei coefficienti (necessaria e sufficiente)

Per il modello di Misura:

- "**Three Measure Rule**": un modello di misura è identificato se ad ogni VL sono associate almeno 3 VM;
- "**Two Measure Rule**": un modello di misura è identificato se ad ogni VL sono associate almeno 2 VM e ogni costrutto è correlato almeno con un altro costrutto.

Funzioni di Discrepanza

Massima Verosimiglianza

$$F_{ML} = \log |\mathbf{C}| + tr \mathbf{S} \mathbf{C}^{-1} - \log |\mathbf{S}| - p + q$$

Minimi Quadrati

$$F_{ULS} = \frac{1}{2} tr \left[\mathbf{S} - \mathbf{C}^2 \right]$$

Minimi Quadrati Generalizzati

$$F_{GLS} = \frac{1}{2} tr \left[\mathbf{S}^{-1} \mathbf{S} - \mathbf{C}^2 \right]$$

Asymptotically Distribution Free

$$F_{ADF/WLS} = \underline{\mathbf{s}} - \underline{\mathbf{c}}^T \mathbf{W}^{-1} \underline{\mathbf{s}} - \underline{\mathbf{c}}$$

Massima Verosimiglianza

Funzione di discrepanza

$$F_{ML} = \log |\mathbf{C}| + tr \mathbf{S} \mathbf{C}^{-1} - \log |\mathbf{S}| - p + q$$

Assunzioni:

→ MULTINORMALITA delle VM

→ S segue una distribuzione di Wishart

→ Sia S che C sono definite positive, quindi esse sono non singolari

Proprietà delle stime:

Gli stimatori di massima verosimiglianza sono asintoticamente corretti, consistenti e asintoticamente efficienti. Inoltre, all'aumentare di N la distribuzione dello stimatore di ML approssima una distribuzione normale.

Minimi Quadrati

Funzione di discrepanza

$$F_{ULS} = \frac{1}{2} tr \left[\mathbf{S} - \mathbf{C}^2 \right]$$

→ Nel caso di campioni di grossa ampiezza gli stimatori ULS forniscono stime molto simili a quelli di massima verosimiglianza, tuttavia le stime ULS non sono asintoticamente efficienti.

→ Al crescere della numerosità campionaria gli ULS forniscono stimatori consistenti senza bisogno di ipotizzare una distribuzione particolare per le VM.

Minimi Quadrati Generalizzati

Funzione di discrepanza

$$F_{GLS} = \frac{1}{2} \text{tr} \left(W^{-1} (S - C)^2 \right)$$

Gli stimatori GLS sono consistenti e all'aumentare della numerosità campionaria la loro distribuzione approssima una normale.

Tuttavia, queste proprietà dipendono dalla scelta di W (caso limite $W=I$ GLS=ULS).

Per garantire le precedenti proprietà agli stimatori GLS la matrice W deve essere scelta sotto l'assunzione che:

1. gli elementi di S siano stimatori corretti delle corrispondenti var/cov

2. gli elementi di S siano asintoticamente distribuiti come una multinormale con media uguale alla corrispondente var/cov e una covarianza asintotica tra s_{ij} e s_{igh} uguale a $N^{-1}(\sigma_{ig}\sigma_{ih} + \sigma_{ih}\sigma_{jg})$

Normalmente  $W = S$

Modelli ad Equazioni Strutturali - corso 2009

Lezione 3

Unidimensionalità dei blocchi riflessivi

Verificare l'unidimensionalità del blocco:

Un modello di misura di tipo **riflessivo** assume che ogni blocco di variabili manifeste sia unidimensionale, in altre parole che un solo concetto latente si rifletta in più indicatori.

La **consistenza interna** di un blocco di variabili manifeste può essere verificata soprattutto attraverso tre indici:

- Sulla base dei risultati di un' **Analisi in Componenti Principali**:
 - Un solo autovalore > 1
 - **Loading Plot (Correlazioni Fattori/Variabili)** per identificare eventuali sottogruppi unidimensionali

- Utilizzando l' **alpha di Cronbach**:
$$\alpha_q = \frac{\sum_{p \neq p'} \text{cor}(x_{pq}, x_{pq'})}{P_q + \sum_{p \neq p'} \text{cor}(x_{pq}, x_{pq'})} \times \frac{P_q}{P_q - 1}$$

- Utilizzando il **rho di Dillon-Goldstein**:
$$\rho_q = \frac{\left(\sum_{p=1}^{P_q} \lambda_{pq} \right)^2}{\left(\sum_{p=1}^{P_q} \lambda_{pq} \right)^2 + \sum_{p=1}^{P_q} \lambda_{pq}^2}$$

Un blocco è considerato unidimensionale se questi ultimi 2 indici sono almeno uguali a 0,7

Il rho di Dillon-Goldstein

Per ogni VM il modello di misura è: $\mathbf{x}_p = \lambda_p^x \xi + \delta_p$

e considerando tutte le VM di un blocco:

$$\sum_p \mathbf{x}_p = \sum_p \lambda_p^x \xi + \sum_p \delta_p$$

Ipotizzando che le VL siano standardizzate, segue:

$$\rho = \frac{\left(\sum_p \lambda_p \right)^2}{\text{Var} \left(\sum_p \mathbf{x}_p \right)} = \frac{\left(\sum_p \lambda_p \right)^2}{\left(\sum_p \lambda_p \right)^2 + \text{Var} \left(\sum_p \delta_p \right)}$$

tutti i
loadings
devono
essere > 0

Unidimensionalità dei blocchi riflessivi

E se il blocco non è unidimensionale?



Le soluzioni possibili:

- **Rimuovere** le variabili manifeste che rendono il blocco non unidimensionale
- Dividere il blocco multidimensionale in **sotto-blocchi unidimensionali**

L'Indice di Comunalità

AVE (Average Variance Extracted) di Fornell & Larcker

La bontà del modello di misura (**reliability of latent variables**) prende in considerazione la parte di variabilità delle VM che la VL riesce a spiegare (**comunalità media**) rispetto alla parte di variabilità

$$AVE_j = \frac{\sum_i \lambda_{ij}^2 \times \text{var } \xi_j}{\sum_i \lambda_{ij}^2 \times \text{var } \xi_j + \sum_i \text{var } \varepsilon_{ij}}$$

Nel caso di VM standardizzate:

$$Communalit\grave{y}_j = \frac{\sum_i \lambda_{ij}^2}{p_j}$$

$$Communalit\grave{y} = \frac{\sum_j p_j \times Communalit\grave{y}_j}{p}$$

é inoltre possibile testare l'ipotesi $H_0: \lambda_{ij}=0$ per ogni loadings

Una VL é considerata un buon predittore delle VM ad essa associate se l'AVE é >0.5 , se l'ipotesi H_0 é rifiutata

e se i loadings standardizzati sono tutti >0.707

Variabilità discriminante delle VL

Le variabili latenti devono essere correlate ma DEVONO anche misurare concetti diversi.

Si può testare l'ipotesi nulla che la correlazione tra due variabili latenti sia più piccola di 1

$$H_0 : \text{cor } \xi_j, \xi_{j'} = 1$$

Regole di decisione:

L'ipotesi nulla è rifiutata se:

- 1) L'intervallo di confidenza al **95%** per ogni correlazione tra le VL non deve comprendere 1;
- 2) $\Delta\chi^2 > 4$ dove $\Delta\chi^2 = \chi^2_{HO} - \chi^2_{FREE}$;
- 3) $(AVE_j \text{ e } AVE_{j'}) > \text{cor}^2 \hat{\xi}_j, \hat{\xi}_{j'}$, in altre parole ogni VL spiega più le proprie VM che le altre VL

La Monofattorialità delle VM

Ogni VM deve essere maggiormente correlata alla propria VL che alle altre VL nel modello:

$$\text{COR } x_{ij}, \xi_j \gg \text{COR } x_{ij}, \xi_{j'}$$

Si può fare riferimento ai Cross-loadings e agli Indici di Modifica per individuare le corrette associazioni tra le VM e le VL

Validazione del modello

Un modello non può essere considerato VERO in assoluto, l'unica cosa possibile è VERIFICARE la sua congruenza con i dati osservati

La VALIDAZIONE di un modello SEM stimato con i metodi covariance-based si basa sulla funzione di DISCREPANZA ed in particolare sull'analisi dei residui del modello, cioè degli scarti tra matrice di var/cov osserva e quella implicata dal modello.

Steps del processo di validazione

- Verifica dell'adeguamento (fit) globale del modello ai dati
- Test statistici sui legami statistici tra le variabili nel modello (siano esse latenti o manifeste)

Indici di validazione globale

Gli indici di validazione globale del modello si basano sulla
FUNZIONE DI DISCREPANZA (fitting function):

$$f(\mathbf{S}, \mathbf{C})$$

Si può dimostrare che:

$$f(\mathbf{S}, \mathbf{C}) \approx \chi_{DF}^2$$

Con $DF =$ gradi di libertà

$$DF = \left\lfloor \frac{1}{2} (P + Q)(P + Q + 1) - t \right\rfloor$$

Indici di validazione globale

Il test del Chi-2

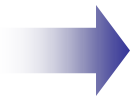
$H_0 : S = C \Leftrightarrow$ Buon adattamento del modello ai dati

$H_1 : S \neq C$

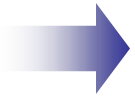
Regola di decisione:

se $p\text{-value} > 0,05$ allora accettiamo $H_0 \Leftrightarrow$ Buon adattamento

Questo test dipende molto dalla numerosità campionaria (N)



Se N è grande, rischio di rifiutare il modello anche con un buon adattamento ai dati!



Difficile procedere al confronto tra il fit di modelli di numerosità diversa

Indici di validazione globale

Il Goodness of Fit Index (GFI)

$$GFI = 1 - \frac{F}{F_{IND}}$$

- Varia tra 0 e 1, ma può succedere che si osservano valori al di fuori di quest'intervallo
- È valido solo per fit function stimate con gli stimatori ML, ULS e GLS

Regola di decisione: un modello è accettato se $GFI > 0,9$

Permette il confronto tra modelli stimati su campioni di numerosità diversa, però NON tiene conto dei gradi di libertà

Indici di validazione globale

L' Adjusted Goodness of Fit Index (AGFI)

$$AGFI = 1 - \frac{1 - GFI}{\frac{DF_{IND}}{DF}}$$

→ Varia tra 0 e 1

→ É valido solo per fit function stimate con gli stimatori ML, ULS e GLS

Regola di decisione: un modello è accettato se $AGFI > 0,9$

Permette il confronto tra modelli stimati su campioni di numerosità diversa, tiene conto dei gradi di libertà, ma NON conosciamo la funzione di distribuzione

Indici di validazione globale

Il Root Mean Squared Residual (RMR)

$$RMR = \sqrt{\frac{1}{k} \sum (s_{ij} - \sigma_{ij})^2}$$

- È una media dei quadrati dei residui
- Non ha un limite superiore
- Non dipende da N
- Può essere usato per confrontare modelli diversi (anche NO-NESTED)

Non si conosce la sua distribuzione statistica

Baseline o Confronto tra Modelli

Modello SATURO:

È quel modello che contiene tanti parametri stimati quanta informazione disponibile. [È il modello con meno vincoli possibili]

Il modello di INDEPENDENZA:

Ha come parametri del modello sole le varianze delle VM. In altre parole è quel modello che assume che non ci sia nessuna relazione tra le variabili del modello. Questo modello ha il max dei gradi di libertà possibili. [È il modello il più restrittivo possibile e qualsiasi test porterà sempre al suo rifiuto]

Indici di Validazione basati sul confronto con un modello baseline

L'indice Comparativo di Bentler (CFI)

Il CFI confronta il modello studiato con il modello corrispondente al caso di **indipendenza** tra le VM:

$$CFI = \frac{\left[n - 1 \quad F_{IND} - DF_{IND} \right] - \left[n - 1 \quad F - DF \right]}{n - 1 \quad F_{IND} - DF_{IND}}$$

dove F_{IND} = è il minimo di F per il modello di indipendenza
e DF_{IND} = gradi di libertà del modello di indipendenza

Un modello è accettato se il CFI > 0.9

Indici di Validazione basati sul confronto con un modello baseline

L'indice Comparativo di Bentler-Bonnet Normalizzato
Bentler-Bonnet Non-Normed Fit Index (NNFI)

Anche questo indice confronta il modello studiato con il modello corrispondente al caso di **indipendenza** tra le VM:

$$NNFI = \frac{\frac{F_{IND}}{DF_{IND}}}{\frac{F}{DF}}$$
$$= \frac{F_{IND}}{DF_{IND}} \cdot \frac{DF}{n-1}$$

Un modello è accettato se il NNFI > 0.9
O addirittura se NNFI > 0.95

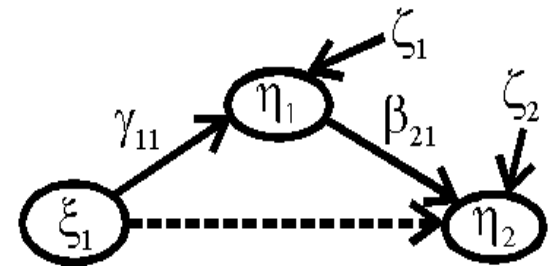
Confronto tra Modelli sulla base di indici basati sull'Information Theory

Tre indici molto simili (tutti basati sull'information theory) possono essere usati per scegliere tra modelli « nested ».

$$AIC = \chi^2 + 2 \text{ n}^\circ \text{ parameters}$$

$$ECVI = \chi^2 / n + 2 \text{ n}^\circ \text{ parameters} / n$$

$$CAIC = \chi^2 + [1 + \ln n] \text{ n}^\circ \text{ parameters}$$



Procedura di Selezione

- 1) Ogni modello è stimato e per ogni modello viene calcolato uno di questi indici.
- 2) Il modello con l'indice più piccolo (che inoltre deve essere più piccolo del corrispondente indice calcolato per il modello saturo) è quello preferito.

Miglioramento del modello

Un volta stabilito che il modello di riferimento si adatta bene ai dati è possibile procedere per passi successivi al MIGLIORAMENTO del modello attraverso l'introduzione di nuovi legami e/o la cancellazione di legami già presi in considerazione.....



Indici di Modifica

Un indice di modifica misura la diminuzione del valore del Chi-2 quando un nuovo legame è aggiunto nel modello.

È inoltre interessante valutare l'aumento del Chi-2 quando un nuovo vincolo è imposto ai parametri del modello, il che può equivalere anche ad « eliminare » un legame, cioè ad imporlo = 0.

Miglioramento del modello

Eliminare un legame dal modello.....

Il test t per la significatività di un coefficiente (N elevato)

$$H_0: \beta = 0$$

$$H_1: \beta \neq 0$$

$$T = \frac{B - 0}{S_B} \approx N$$

Regola di decisione:

Se $|T|_{S_B} > 2$ rifiuto H_0

Miglioramento del modello

Aggiungere un nuovo legame al modello.....

Per ogni parametro fisso viene calcolato di quanto aumenterebbe la statistica test del Chi-2 se quel parametro fosse incluso tra i parametri da stimare

L'indice di modifica è il **rapporto tra il Chi-2 del modello con quel parametro fisso ed il modello con quel parametro libero di essere stimato** → **si distribuisce come un Chi-2 con 1 solo grado di libertà**

Un parametro fisso è incluso nel modello se la statistica test corrispondente è > 4 .

Tuttavia la **respecificazione** del modello deve essere sempre accompagnata alla « giustificazione » teorica di quel nuovo legame (ottenere risultati sorprendenti è sempre interessante....ma solo se i nuovi legami aggiunti sono coerenti con la teoria che ha definito il modello);

Test sui Path Coefficients

Un legame strutturale deve essere eliminato dal modello se il **coefficiente di regressione** ad esso associato non è statisticamente significativo

AMOS calcola per le stime non standardizzate di ogni coefficiente di regressione il suo standard error (**s**), in seguito la stima del coefficiente di regressione è divisa per lo standard error (**C.R.** for **Critical Ratio**) e calcola il p-value (**p**) associato con l'ipotesi nulla che questo coefficiente sia 0.

Regola di decisione:

• Coefficienti con un **C.R.** > 2 sono considerati significativamente diversi da zero e quindi il corrispondente legame è mantenuto nel modello.

• Coefficienti con un **C.R.** < 2 sono considerati non significativamente diversi da zero e quindi il corrispondente legame è eliminato dal modello.

Analisi dei Residui

Per ogni coppia di VM (x_i, x_k), il residuo è la differenza tra il valore della cov osservato e quello stimato attraverso il modello. Gli errori standard dei residui sono stimati così che (asintoticamente) i **residui standardizzati** (cioè indipendenti dalle unità di misura delle variabili) siano:

$$e_{ik}^* = \frac{S_{ik} - C_{ik}}{\hat{\sigma}_{C_{ik}}}$$

Interpretazione:

Un **residuo positivo** > 2 indica che la covarianza tra le 2 VM è **sottostimata** (bisogna aggiungere altri legami).

Un **residuo negativo** < -2 indica che la covarianza tra le 2 VM è **sovrastimata** (bisogna rimuovere dei legami che impattano su queste variabili).

Approfondimenti (1)

- L'unità costitutiva di un Modello ad Equazioni Strutturali è la *retta di regressione* detta *equazione strutturale*: essa esprime una relazione causale tra una variabile dipendente e una o più variabili indipendenti
- Considerare solo i nessi causali tra le variabili dipendenti e indipendenti rende banale e semplicistica la rappresentazione della realtà, dove esiste una interazione anche tra le variabili indipendenti
- Il passaggio da una singola equazione di regressione ad un sistema di equazioni, comporta l'adozione di un diverso metodo per la stima dei parametri del modello, che nel caso della regressione avviene, generalmente, attraverso il metodo OLS
- Nel caso del sistema di equazioni la proprietà di indipendenza tra gli errori e le variabili indipendenti non è detto che sia verificata

Approfondimenti (2)

- Il punto di partenza di LISREL è la matrice di *varianza-covarianza* fra le variabili osservate
- L'obiettivo è la stima dei parametri del modello che esprimono i nessi causali fra le variabili
- Il problema in cui si incorre è che una stessa matrice di *varianza-covarianza* può essere generata da differenti modelli causali, ma un modello causale può produrre una sola matrice di *varianza-covarianza*
- Pertanto è possibile determinare una matrice *varianza-covarianza* teorica che sia il più possibile simile a quella osservata
- È necessario definire un certo numero di parametri che costituiranno le incognite del modello da stimare: si cercheranno i valori dei parametri che, sostituiti nel modello, generano la matrice di *varianza-covarianza* con lo scarto minore da quella osservata

Approfondimenti (3)

Le fasi di LISREL sono quattro:

- *Formulazione del modello*, in modo che sia matematicamente risolvibile ovvero identificato
- *Stima dei parametri strutturali*, attraverso un processo iterativo che minimizza la differenza tra l'osservato e il teorico
- *Verifica del modello*, in cui si confronta il modello teorico con i dati osservati per la falsificazione
- *Modifica del modello*, nel caso in cui il modello è falsificato esso va modificato e stimato nuovamente

Il punto di partenza su cui si itera l'intera procedura è la matrice di *varianza-covarianza* tra le variabili X e Y, esprimibile in funzione delle 8 matrici del modello, per cui cambiando il modello teorico, cambia la matrice attesa (prodotta dal modello)

La matrice Σ attesa implicata dal modello non potrà mai essere uguale alla matrice S osservata a causa della variabilità campionaria e dell'eventuale mal specificazione del modello: il residuo della differenza sarà usato per la falsificazione del modello

Approfondimenti (4)

Le fasi di LISREL sono quattro:

- *Formulazione del modello*, in modo che sia matematicamente risolvibile ovvero identificato
- *Stima dei parametri strutturali*, attraverso un processo iterativo che minimizza la differenza tra l'osservato e il teorico
- *Verifica del modello*, in cui si confronta il modello teorico con i dati osservati per la falsificazione
- *Modifica del modello*, nel caso in cui il modello è falsificato esso va modificato e stimato nuovamente

Il punto di partenza su cui si itera l'intera procedura è la matrice di *varianza-covarianza* tra le variabili X e Y, esprimibile in funzione delle 8 matrici del modello, per cui cambiando il modello teorico, cambia la matrice attesa (prodotta dal modello)

La matrice Σ attesa implicata dal modello non potrà mai essere uguale alla matrice S osservata a causa della variabilità campionaria e dell'eventuale mal specificazione del modello: il residuo della differenza sarà usato per la falsificazione del modello

Approfondimenti (5)

Il metodo usato per la stima dei parametri è la Massima Verosimiglianza.

- Le 8 matrici di LISREL contengono dei parametri fissi, valori assegnati imm modificabili e dei parametri liberi cioè da stimare
- La massima varosimiglianza permette di *stimare i parametri incogniti della popolazione individuando quei parametri che generano la più elevata probabilità per i dati campionari di essere osservati*
- Quindi, data una certa matrice di covarianza osservata (campione) S , dobbiamo individuare la probabilità che questa matrice derivi da una matrice teorica Σ (popolazione)
- Per calcolare tale probabilità useremo la distribuzione di Wishart, espressa in funzione dei parametri del modello

Approfondimenti (6)

I parametri del modello saranno stimati in modo da massimizzare la probabilità che S derivi da Σ , ovvero i parametri che massimizzano la distribuzione di Wishart: quella che cambia è Σ

Per trovare il massimo si studiano le derivate e le derivate parziali, rispetto ai parametri delle equazioni componenti il sistema

Essendo liberi ed interdipendenti, si cercheranno i parametri tale che la loro combinazioni massimizzi la funzione di Wishart: da qui l'iteratività della procedura di stima

Il processo, tuttavia, non si innesta con valori arbitrari per i parametri, ma essi sono stimati con il metodo dei minimi quadrati a due stadi: su tali stime si innesta il metodo della massima verosimiglianza

La mtrice attesa è modificata fino a che non is raggiunge il minimo della funzione: se un ulteriore modifica dei paramtri non produce miglioramenti nella differenza tra l'osservato e l'atteso, l'algorithmo si arresta

Se la differenza non è accettabile si rifiuta il modello