

多変量解析法の分類

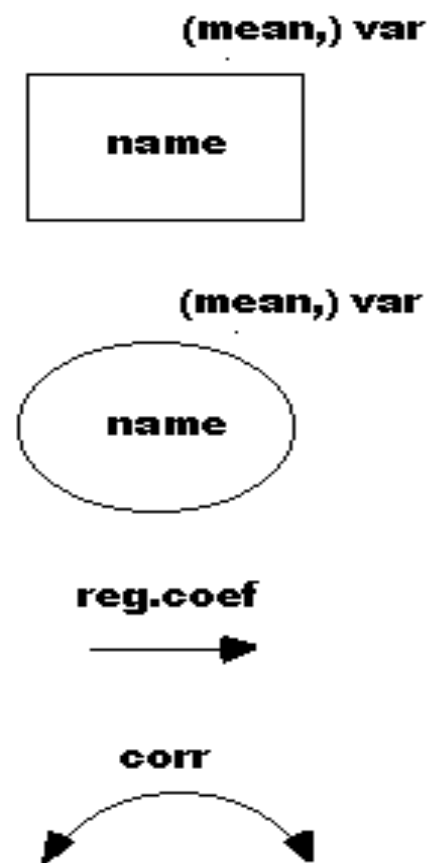
表 1: 多変量解析法の分類

		説明変数 X は観測されているか	
		観測されている (顕在変数)	観測されていない (潜在変数)
説明変数 X は連続か離散か	連続	回帰分析 冗長性分析 (判別分析) (ロジスティック回帰)	因子分析 主成分分析 (IRT) (カテゴリカルデータの因子分析) (数量化 III 類, CA)
	離散	分散分析 (対数線形モデル)	クラスタ分析 (潜在クラスモデル)

表中、括弧の中は、基準変数 Y がカテゴリカルな場合、それ以外は、基準変数が連続の場合を表す。また、説明変数が観測されている場合、その連続離散の区別はあまり重要ではない。

パス図 (Path Diagram)

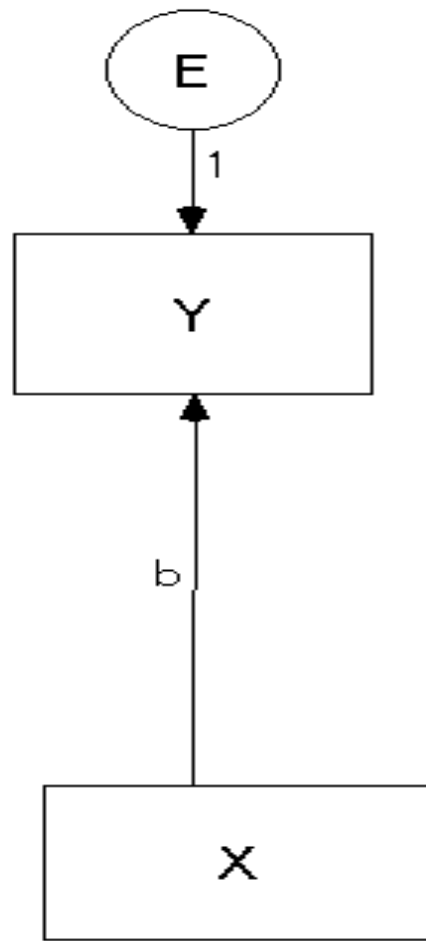
- 四角は観測される変数
(変数名を中に、平均分散等を肩)
- (楕)円は観測されない変数
(潜在変数)
- 矢印は因果の流れ
(回帰係数を肩に)
- 曲がった矢印は相関関係
(相関係数を肩に)



パス図の読み方

- 外性変数（独立変数）
入ってくる矢印がない。
- 内性変数（従属変数）
少なくともひとつ入ってくる矢印がある。
（内性変数には誤差項を付ける。）
- 入ってくるすべての矢印に係数を掛けて加える。

回歸分析

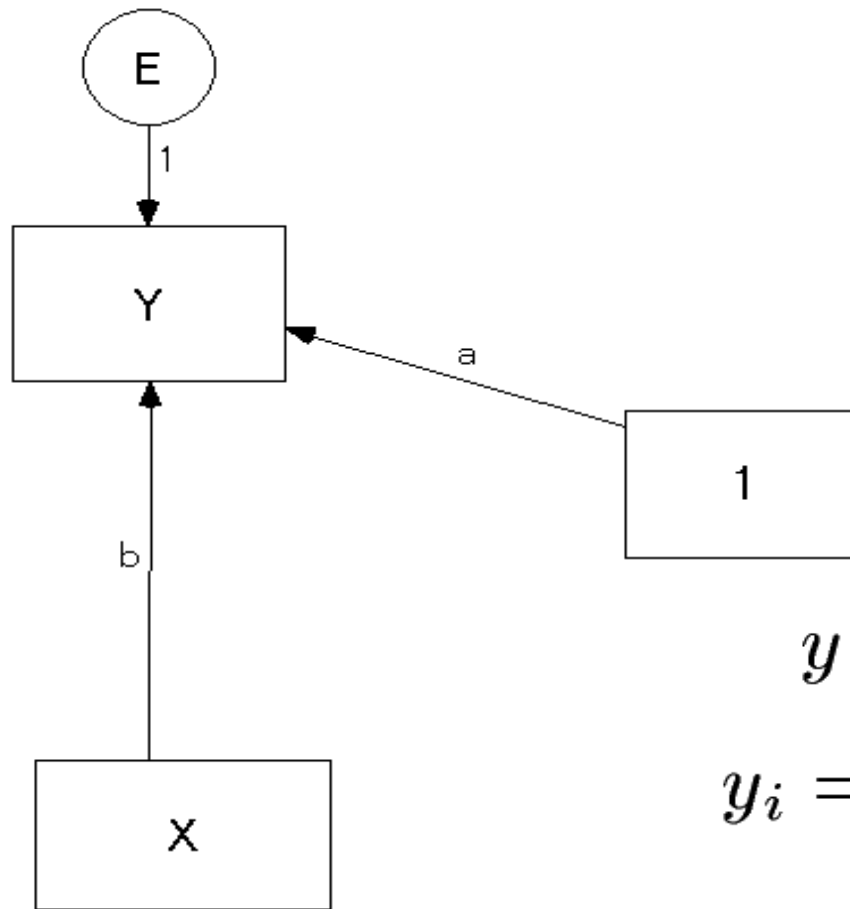


$$y = \beta x + e$$

$$y_i = \beta x_i + e_i$$

$$i = 1, 2, \dots, N$$

回帰分析 (切片付き)



$$y = \alpha + \beta x + e$$

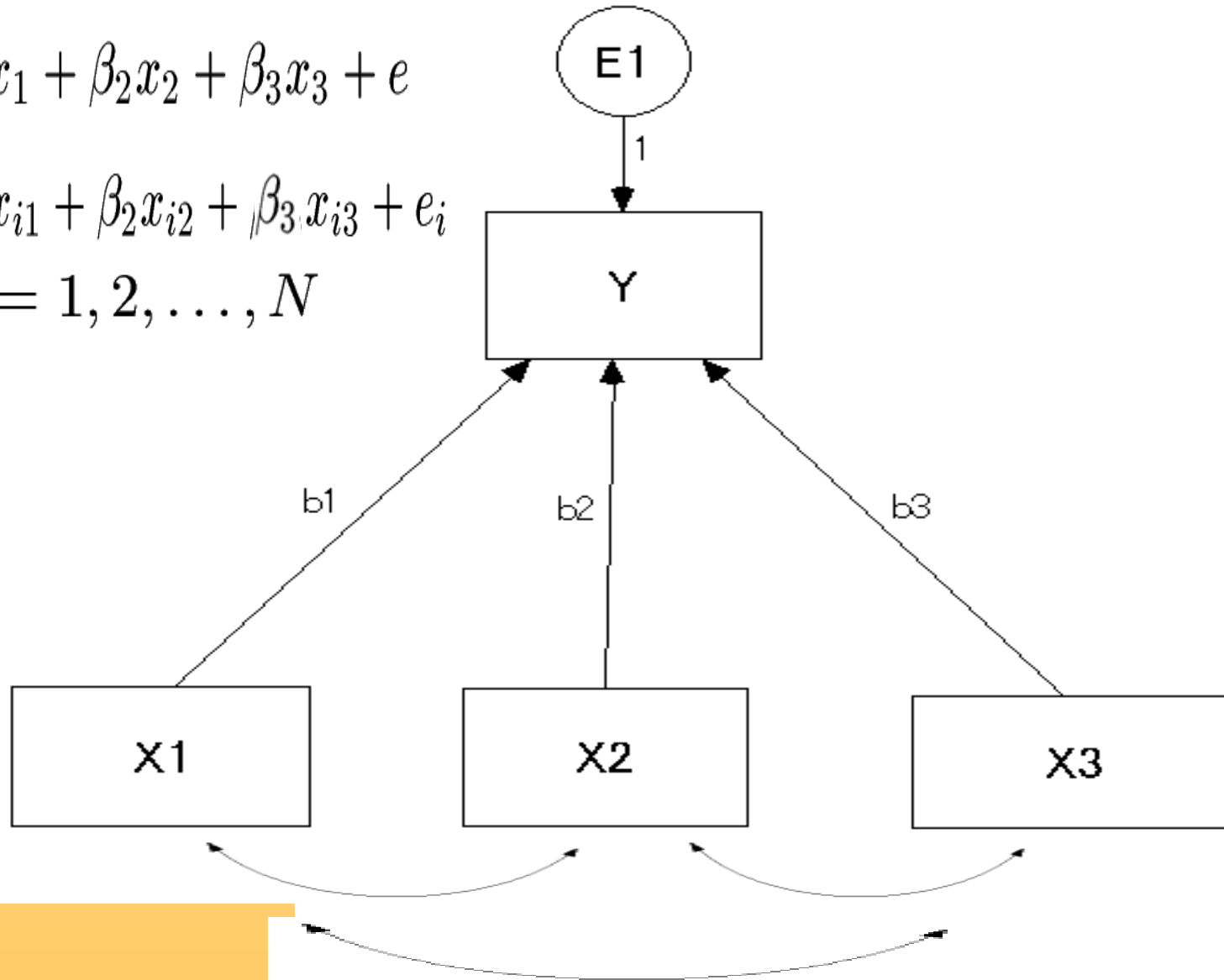
$$y_i = \alpha + \beta x_i + e_i$$

$$i = 1, 2, \dots, N$$

重回歸分析

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + e$$

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i$$
$$i = 1, 2, \dots, N$$

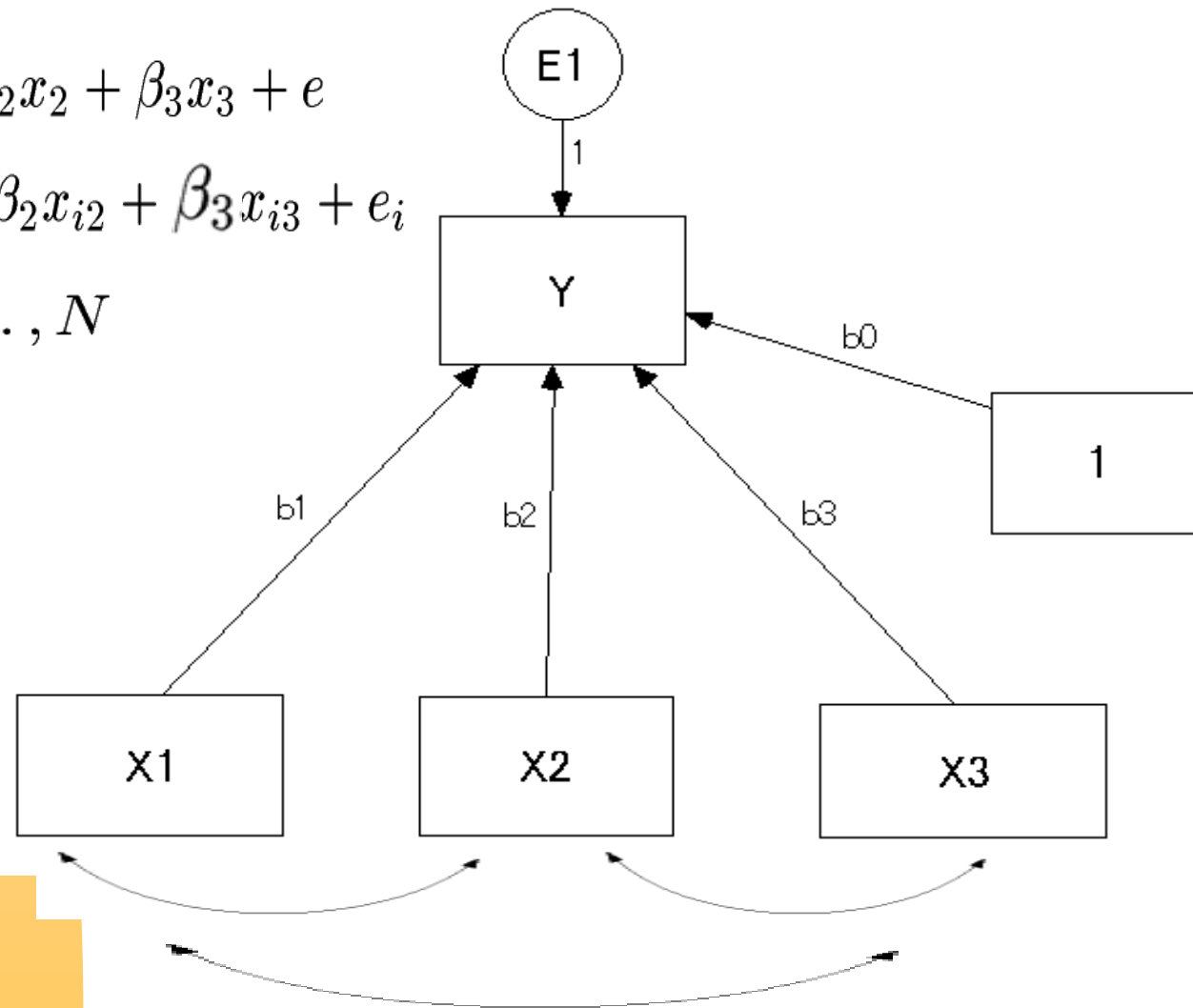


重回帰分析（切片付き）

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + e$$

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i$$

$$i = 1, 2, \dots, N$$



重回帰分析（数式）

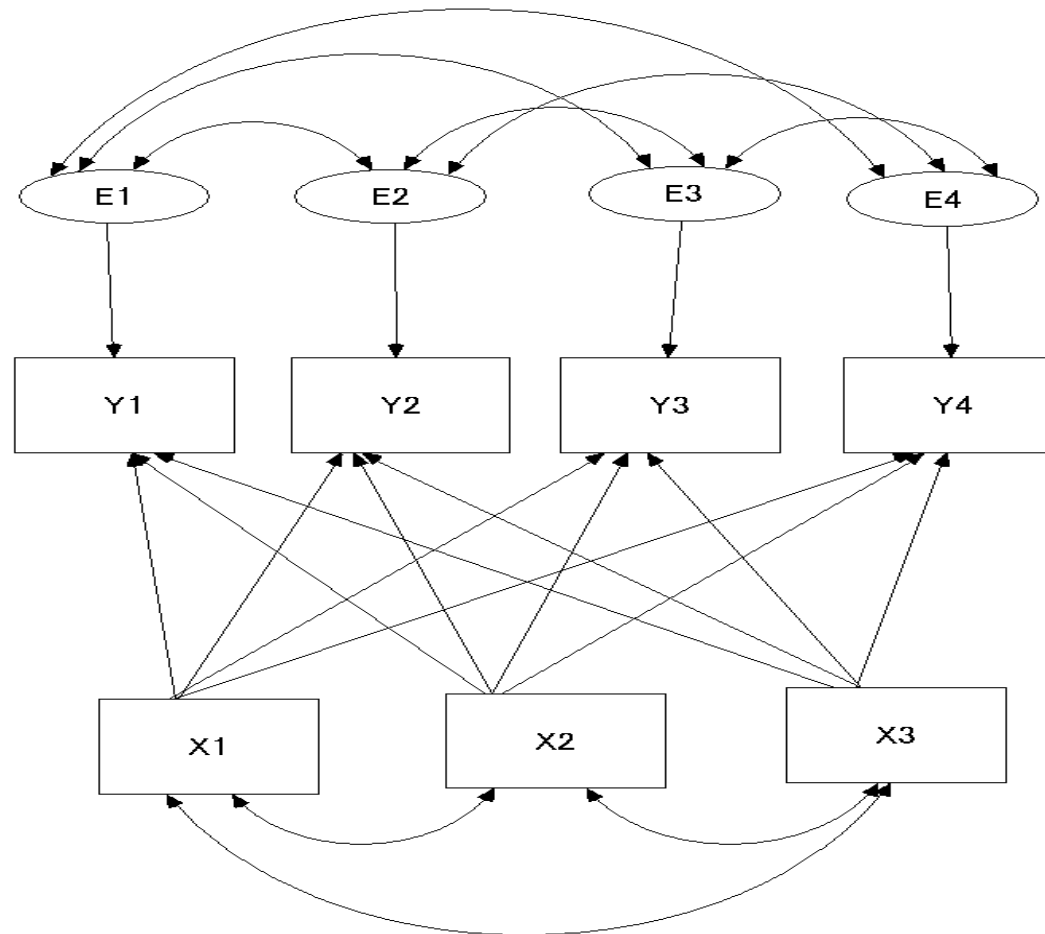
$$y_i = \beta_0 + \sum_{j=1}^q \beta_j x_{ij} + e_i$$

$i = 1, 2, \dots, N$

$x_0 = \mathbf{1}$ とすると

$$y_i = \sum_{j=0}^q \beta_j x_{ij} + e_i$$

多变量（重）回归分析



多變量 (重) 回歸分析 (數式)

$$y_{ij} = \beta_{j0} + \sum_{k=1}^s \beta_{jk} x_{ik} + e_{ij}$$

回帰分析のベクトル表示

行列で書くと

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$$

ただし、

$$\mathbf{y} = N \times 1, \quad \mathbf{X} = N \times (q + 1),$$

$$\mathbf{e} = N \times 1$$

回帰係数の推定

- 最小自乗法

$$\text{RSS} = \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

q

$$\hat{y}_i = \sum_{j=0} \beta_j x_{ij}$$

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$$

回帰係数の推定の詳細

$$\begin{aligned}\frac{\partial \text{RSS}}{\partial \beta_k} &= 2 \sum_{i=1}^n (y_i - \hat{y}_i) \times \frac{\partial (y_i - \hat{y}_i)}{\partial \beta_k} \\ &= 2 \sum_{i=1}^n (y_i - \hat{y}_i) \times \frac{\partial \sum_{j=0}^q x_{ij} \beta_j}{\partial \beta_k} \\ &= 2 \sum_{i=1}^n (y_i - \hat{y}_i) \times x_{ik} \\ &= 2 \sum_{i=1}^n (y_i - \sum_{j=0}^q x_{ij} \beta_j) \times x_{ik} \\ &= 2 \sum_{i=1}^n y_i x_{ik} - 2 \sum_{i=1}^n \left(\sum_{j=0}^q x_{ij} \beta_j \right) x_{ik} \\ &= 2 \sum_{i=1}^n y_i x_{ik} - 2 \sum_{i=1}^n \left(\sum_{j=0}^q x_{ij} x_{ik} \beta_j \right) \\ &= 2 \sum_{i=1}^n y_i x_{ik} - 2 \sum_{j=0}^q \left(\sum_{i=1}^n x_{ij} x_{ik} \right) \beta_j\end{aligned}$$

$\sum_{i=1}^n y_i x_{ik}$ は $X' \mathbf{y}$ の第 k 要素、 $\sum_{i=1}^n x_{ij} x_{ik}$ は $X' X$ の第 jk 要素したがって、方程式 $\frac{\partial \text{RSS}}{\partial \beta_k} = 0, k = 0, 1, 2, \dots, q$ は、まとめて

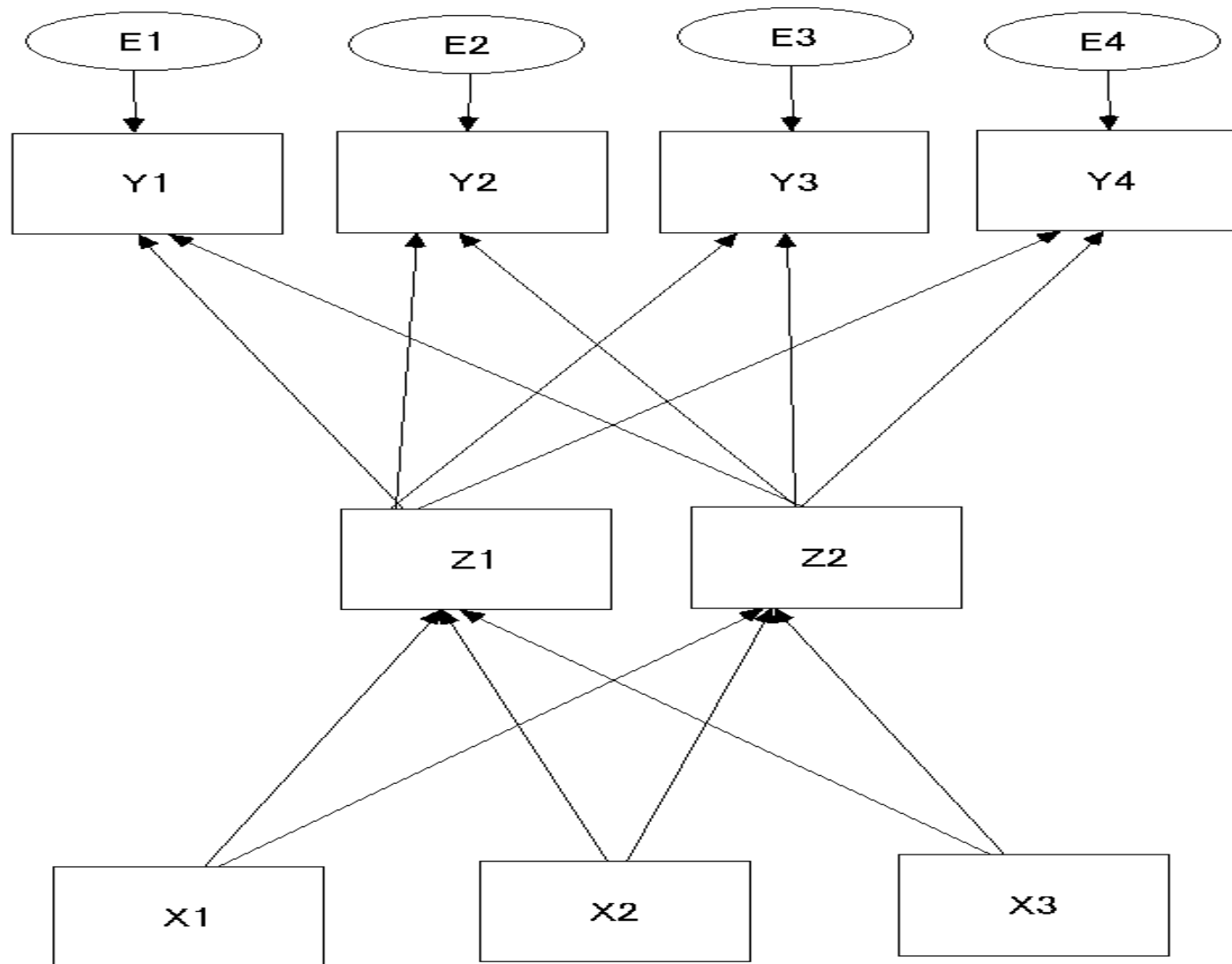
$$X' \mathbf{y} - X' X \boldsymbol{\beta} = \mathbf{0}$$

回帰係数の推定（単回帰）

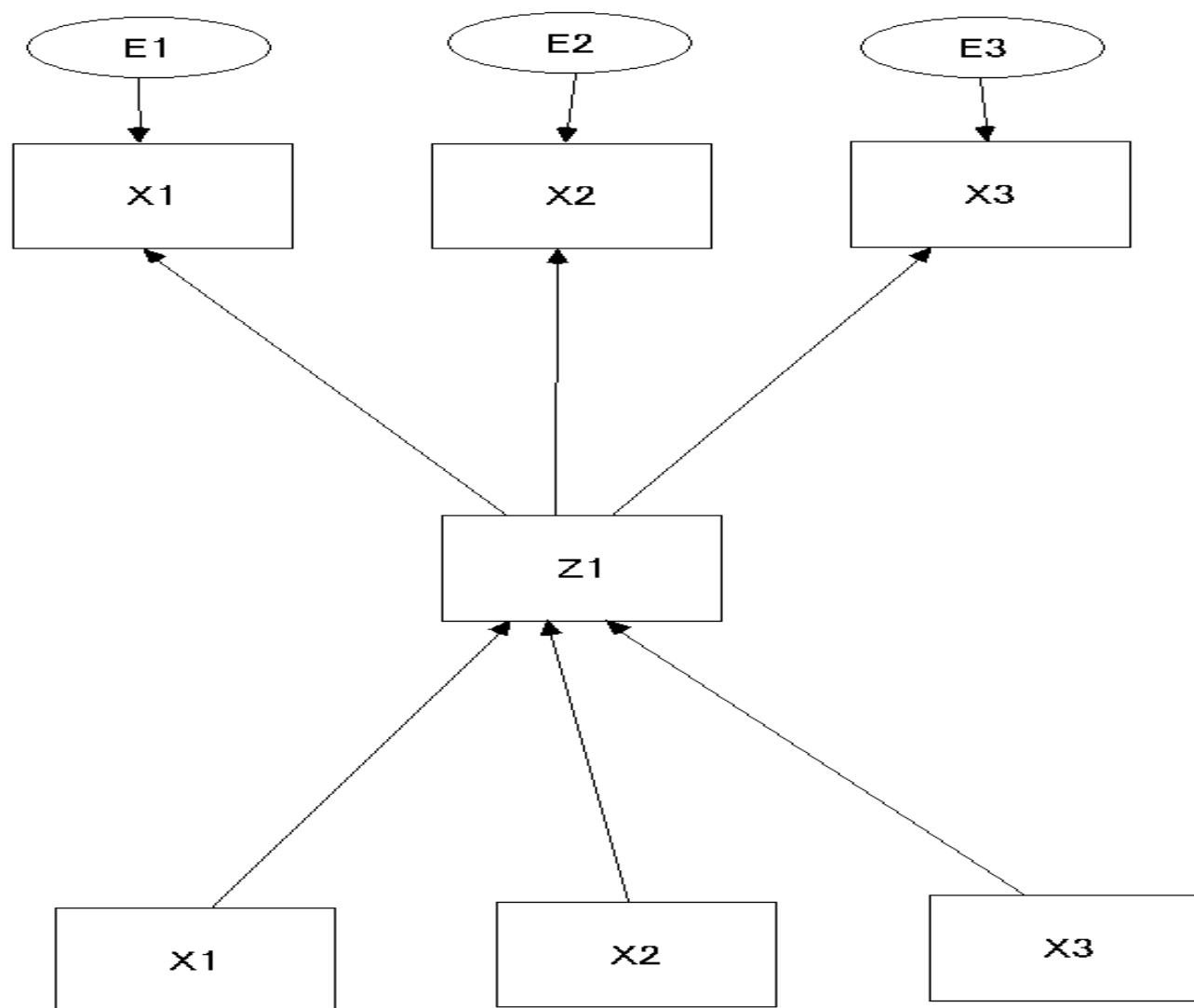
$$\hat{\beta} = \frac{\text{cov}(x, y)}{\text{var}(x)}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

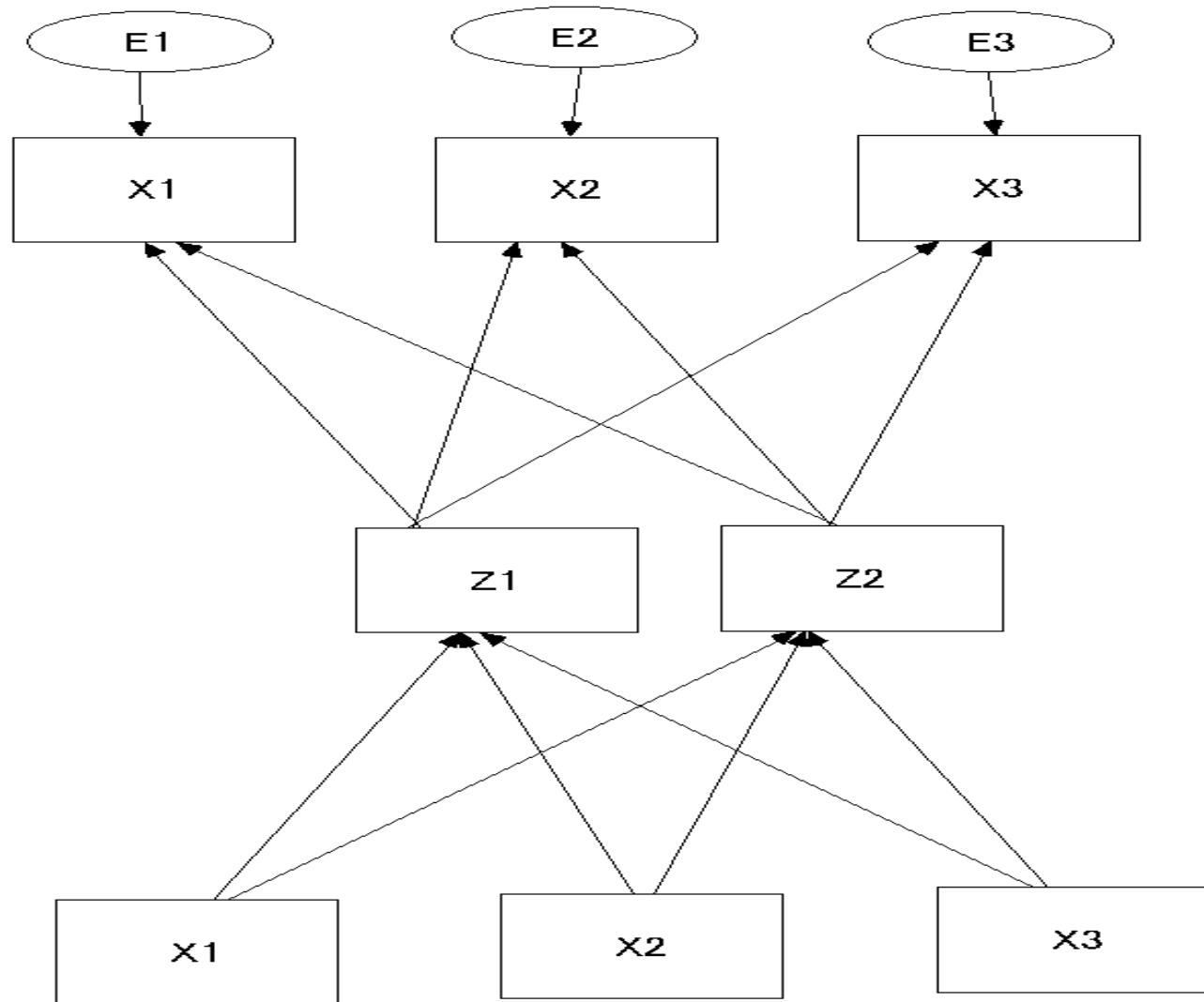
冗長性分析



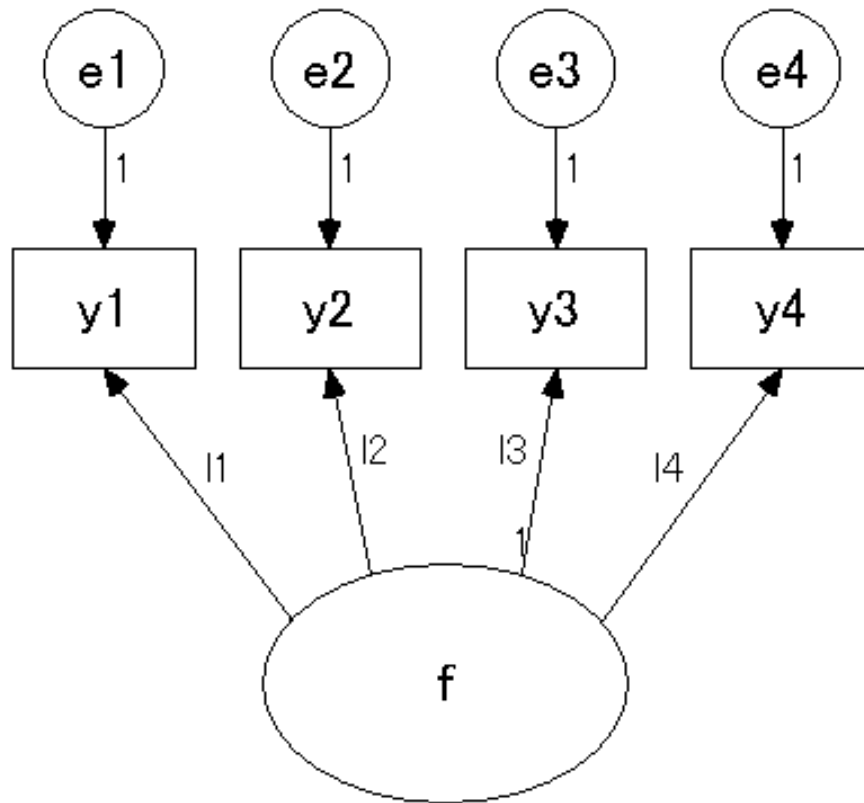
主成分分析 (1 因子)



主成分分析

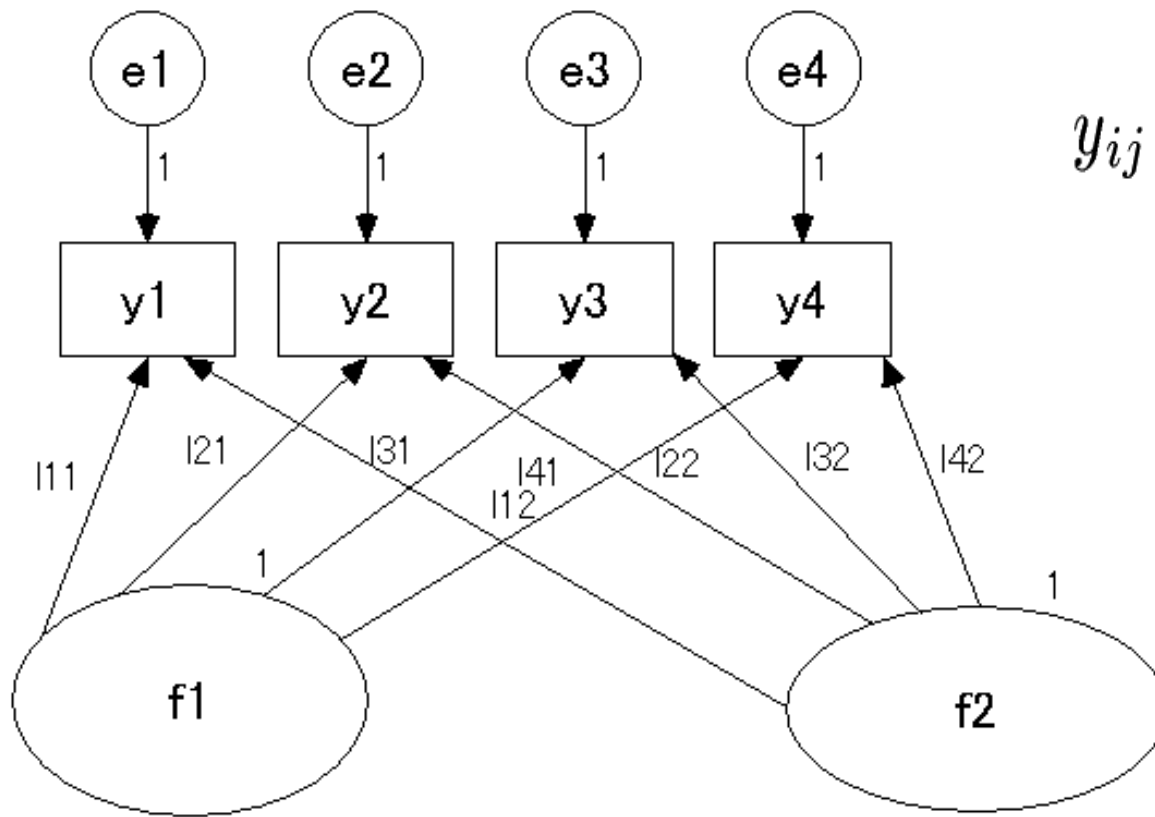


因子分析（1因子モデル）



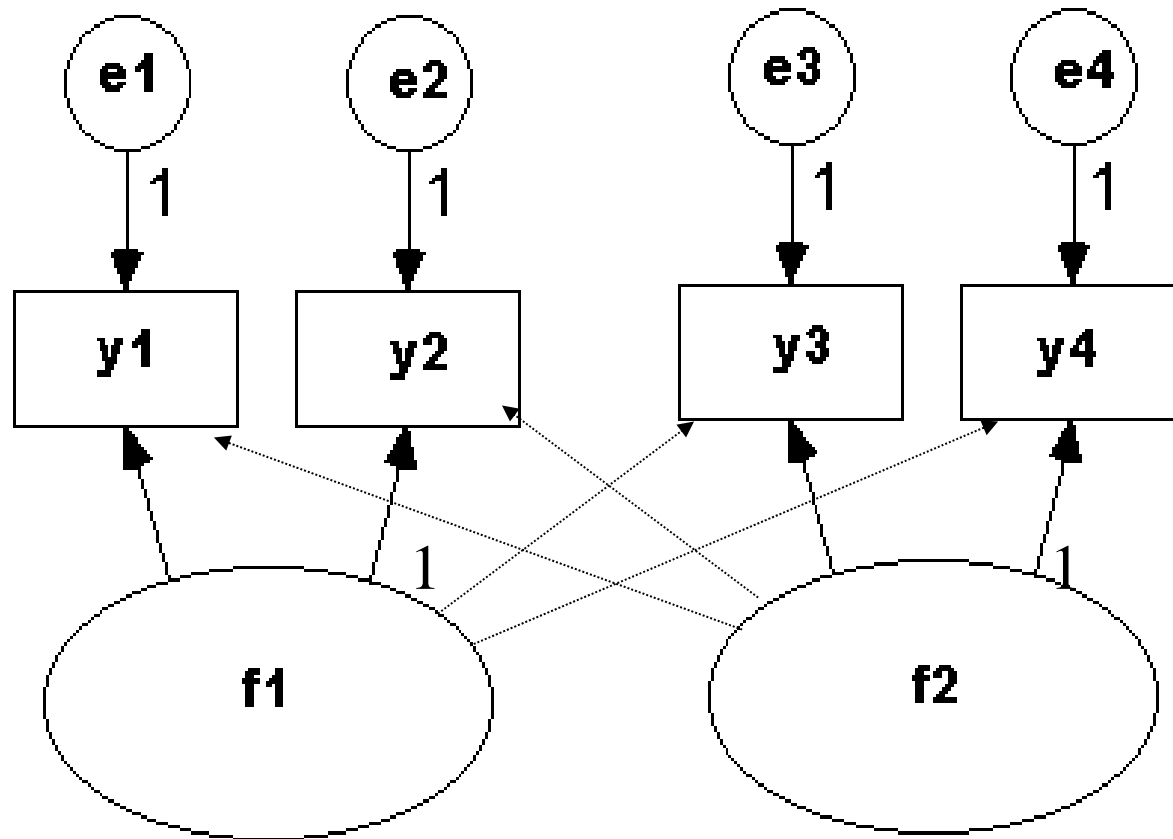
$$y_{ij} = f_i \lambda_j + e_{ij}$$

因子分析（多因子モデル）

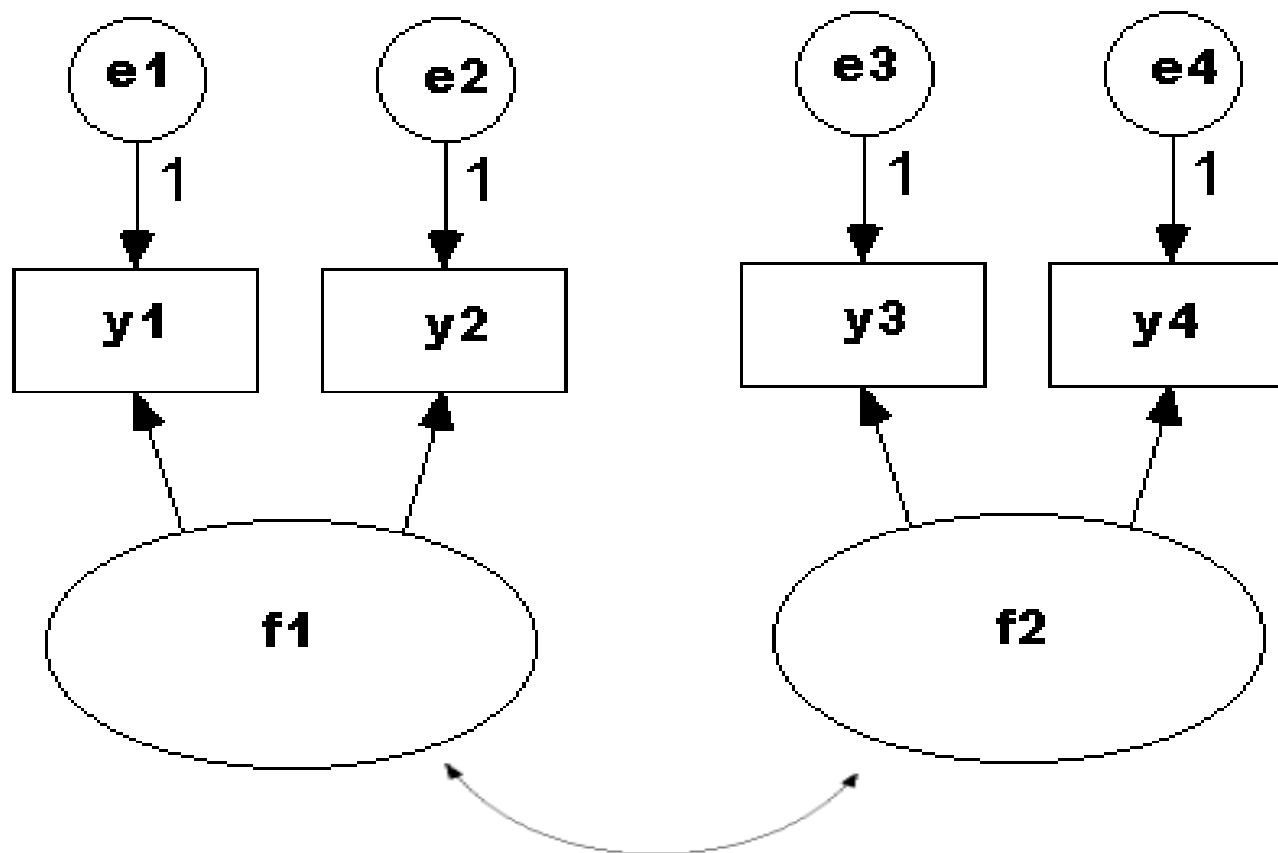


$$y_{ij} = \sum_{q=1}^r f_{iq} \lambda_{jq} + e_{ij}$$

単純構造 (回転)

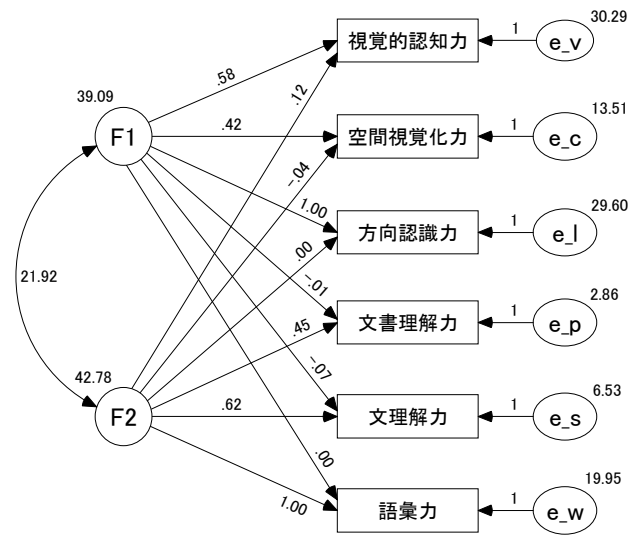


斜交因子



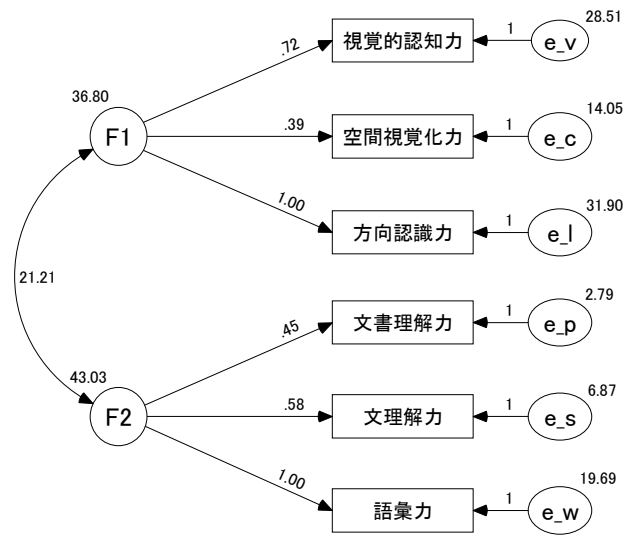
相関係数

Chi-square = 1.071 (4 df)
p = .899



Example 20: Model 2
Two unconstrained factors
Holzinger and Swineford (1939) data
非標準化推定値

Chi-square = 3.638 (8 df)
p = .888



Example 20: Model 2R
Restricted two-factor model
Holzinger and Swineford (1939) data
非標準化推定値

AIC=29.231
GFI=.951

