1時間でだいたいわかる

構造方程式モデリング

Structural Equation Modeling (SEM)

構造方程式モデリングとは何か

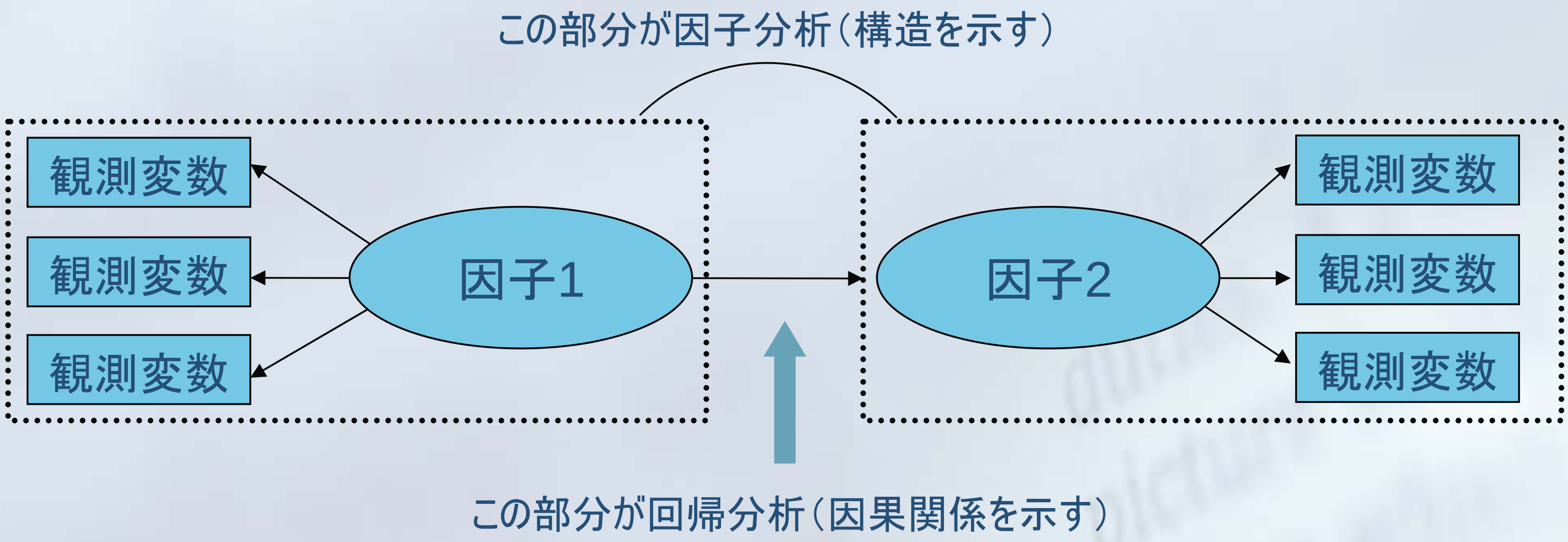
- 構造方程式モデリング（Structural Equation Modeling, SEM）とは：
  - 別名、共分散構造分析（covariance structural analysis）
  - 構成概念や観測変数の性質を調べるために集めた多くの観測変数を同時に分析するための統計的方法
- 本来、構造方程式モデリングは主に以下の3つを含みます。
  - 共分散構造分析（covariance structural analysis）
  - 潜在混合分布モデル（latent mixture model)
  - 潜在クラスモデル（latent class model）

モデル	観測変数		潜在変数		潜在変数のない多変量解析との対応
	連続変数	質的変数	連続変数	質的変数	
共分散構造分析	○	△	○		回帰分析、判別分析
潜在混合モデル	○			○	数量化I類
潜在クラスモデル		○		○	数量化II類

- 当面、観測変数・潜在変数共に連続変数を用いる共分散構造分析だけを扱いそれをSEMと呼ぶことにします。

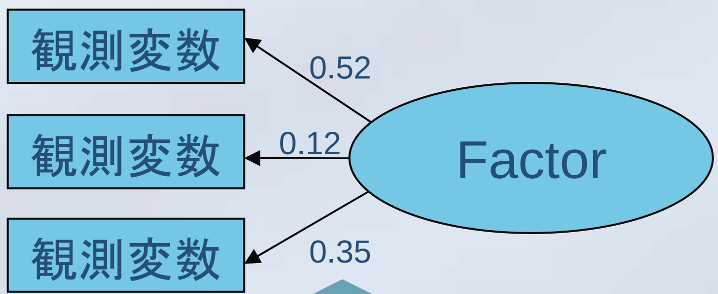
SEMとは何か

- 以下の二つの方程式の合体と言えます。
  - 測定方程式・・・いわゆる因子分析
  - 構造方程式・・・いわゆるパス回帰
- 因子の回帰分析だ！と覚えればわかりやすいと思います。



因子分析ってなんでしたっけ

- つまり、観測された変数は、何らかの隠された要因(Factor)が基になっているという考えで、その要因の影響を「相関」(つまり分散・共分散)で判別しようとうする考え。

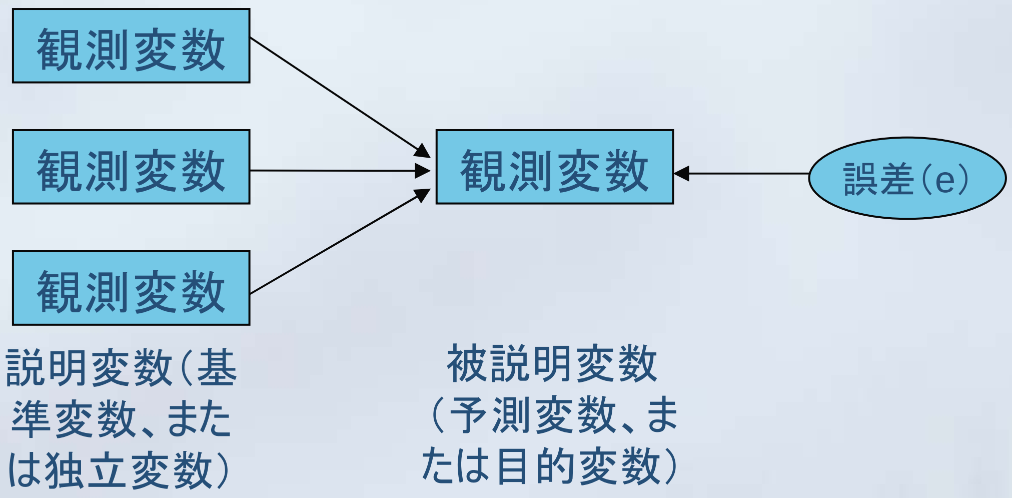


頭文字を取って一般に*f*であらわす。実は、*f*がどのような値を取ろうと、あまり意味はない

Factorがどれだけ影響しているかが重要。この値を因子負荷量という。

回帰分析ってなんでしたっけ

- もちろん簡単なことですが、



- 誤差って何？と思うかも。つい忘れがちです。式で書けば、

$$Y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + e$$

回帰分析の解き方教室

- 教科書によると、回帰分析の母数の推定方法は以下の3通りで、そのいずれでも解は一致します。
  - 最小２乗法
  - 最尤法
  - モーメント法

- ここでは共分散構造分析の基礎となるモーメント法を紹介します。

モーメント法による単回帰モデルの母数推定

- 以下の単回帰式の母数を推定する

$$x_2 = \alpha x_1 + \beta + e$$
$$E[e] = \mu_e = 0$$
$$E[ex_1] = 0$$

*この仮定はつまり、誤差の平均の期待値は0で、誤差と独立変数は無相関であるとしている。これは回帰分析における基本的な仮定。

- 両辺の期待値を計算する

$$E[x_2] = E[\alpha x_1 + \beta + e] = \alpha E[x_1] + E[\beta] + E[e]$$
$$\mu_2 = \alpha \mu_1 + \beta$$
$$\beta = \mu_2 - \alpha \mu_1$$

- 単回帰式の両辺に確率変数 x_1 をかけ、期待値を取る

$$E[x_2 x_1] = \alpha E[x_1 x_1] + \beta E[x_1] + E[ex_1]$$
$$\sigma_{m21} = \alpha \sigma^2{}_{m1} + \beta \mu_1$$
$\beta = \mu_2 - \alpha \mu_1$だから
$$\sigma_{m21} = \alpha \sigma^2{}_{m1} + \mu_1(\mu_2 - \alpha \mu_1) = \alpha \sigma^2{}_{m1} + \mu_2 \mu_1 - \alpha \mu_1 \mu_1$$
$\mathbf{\Sigma} = \mathbf{\Sigma}_m - \boldsymbol{\mu}\boldsymbol{\mu}'$（共分散行列は積率行列から平均の二乗の行列をひいたもの）から
$$\sigma_{m21} - \mu_2 \mu_1 = \alpha(\sigma^2{}_{m1} - \mu_1 \mu_1)$$
$$\sigma_{21} = \alpha \sigma^2{}_1$$
$$\alpha = \sigma_{21} / \sigma^2{}_1$$

続･モーメント法による単回帰モデルの母数推定

- 前のページではごちゃごちゃやりましたが、要は、最終的に以下のようになりました。

$$\alpha = \sigma_{21} / {\sigma^2}_1$$

$$\beta = \mu_2 - \alpha\mu_1$$

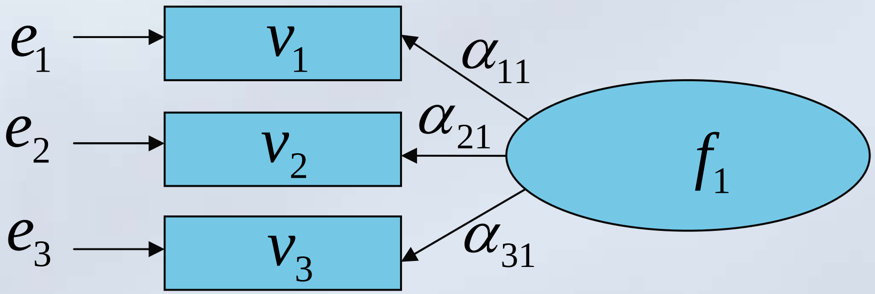
$$\hat{\alpha} = s_{21} / {s^2}_1$$

$$\beta = \bar{x}_2 - \hat{\alpha}\bar{x}_1$$

- ここで大事なのは、母数（パラメータ）のαを、変数の分散と共分散の統計量で推定することができたということです。

さて、測定方程式

- 測定方程式は、前に言ったように、因子分析のことです。別の言い方をすれば、「構成概念」を扱う方程式です。
- 例えば以下のパス図は右の式で表します（変数の添え字は、矢印のさささる方→指す方の順。そうすると行列で都合がいい。）



$$v_1 = \alpha_{11} f_1 + e_1$$
$$v_2 = \alpha_{21} f_1 + e_2$$
$$v_3 = \alpha_{31} f_1 + e_3$$
$$E\left[e_i e_j\right] = 0 (i \neq j)$$
$$E\left[f_i\right] = 0, V\left[f_i\right] = 1$$
$$E\left[e_i\right] = 0, V\left[e_i\right] = \sigma^2{}_{ei}$$
$$E\left[f_i e_j\right] = 0$$

測定方程式の共分散構造

- 共分散を母数の関数で表現することを「構造化」といい、共分散(行列)を方程式モデルの母数で表現したものを「共分散構造」といいます。
- 前ページの測定方程式の場合、以下のようになります。

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & & \\ \sigma_{21} & \sigma_2^2 & \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix} = \begin{bmatrix} \alpha_{11}^2 + \sigma_{e1}^2 & & \\ \alpha_{21}\alpha_{11} & \alpha_{21}^2 + \sigma_{e2}^2 & \\ \alpha_{31}\alpha_{11} & \alpha_{31}\alpha_{21} & \alpha_{31}^2 + \sigma_{e3}^2 \end{bmatrix}$$

- これは、回帰分析のモーメント法と同じことになりました。測定方程式ではfが平均0、分散1に仮定されているので途中の計算で消え、最後は母数だけになってしまうのです。

測定方程式の行列表記

- 測定方程式を行列表記すると以下のようになります。

$$\boldsymbol{v} = \boldsymbol{A}\boldsymbol{f} + \boldsymbol{e}$$

$$E[\boldsymbol{f}] = \boldsymbol{o}$$

$$E[\boldsymbol{e}] = \boldsymbol{o}$$

$$E[\boldsymbol{f}\boldsymbol{e}] = \boldsymbol{O}$$

- そして共分散構造は以下のようになります

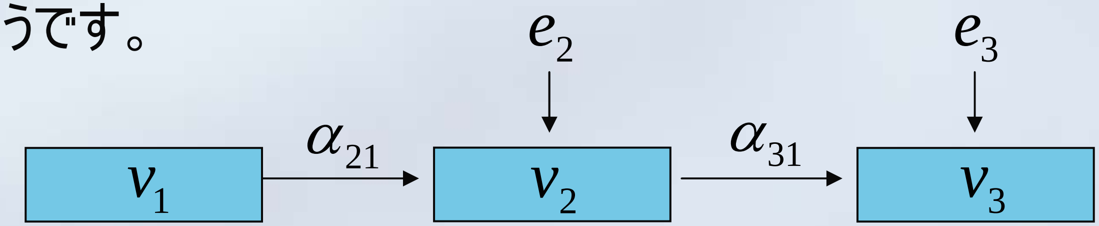
$$\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{\Sigma}_{rf}\boldsymbol{A}' + \boldsymbol{\Sigma}_e$$

潜在変数の相関がある場合にパラメータが含まれます

誤差間に相関がある場合にパラメータが含まれます

構造方程式

- 構造方程式は回帰分析をつないでいくと思えばいい。
- 矢印がさささる変数を内生変数、さらない変数を外生変数といいます。
- 内生変数にはかならず誤差があります。
- 実際のデータでは、無理に潜在変数を作らず、構造方程式を使ったほうがいい場合が多いようです。



$\boldsymbol{v} = \boldsymbol{A}\boldsymbol{v} + \boldsymbol{e}$

$\boldsymbol{e}$はv_iが外生変数であればv_iを、

内生変数であればe_iをi番目の要素として持つ

「残差ベクトル」または「外生変数ベクトル」

$\boldsymbol{v}$は「構造変数ベクトル」

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ \alpha_{21} & 0 & 0 \\ 0 & \alpha_{32} & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ e_2 \\ e_3 \end{bmatrix}$$

構造方程式の共分散構造

- 補足1：残差とは他の構造変数から説明されなかった残りであるから、他から説明されなかった変数（外生変数）は、その変数自身が残差となる
- 補足2：$\boldsymbol{A}$ の対角成分は常に0
- 補足3：v_i が外生変数であれば $\boldsymbol{A}$ の i 行は常にゼロベクトル $\boldsymbol{o}'$
- 共分散構造の行列表記

$$\boldsymbol{v} = \boldsymbol{A}\boldsymbol{v} + \boldsymbol{e}$$

$$\boldsymbol{I}\boldsymbol{v} = \boldsymbol{A}\boldsymbol{v} + \boldsymbol{e}$$

$$(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{v} = \boldsymbol{e}$$

$$\boldsymbol{T} = (\boldsymbol{I} - \boldsymbol{A})^{-1}$$

$$\boldsymbol{v} = \boldsymbol{T}\boldsymbol{e}$$

$$\boldsymbol{\Sigma} = \boldsymbol{T}\boldsymbol{\Sigma}_e\boldsymbol{T}'$$

逆に言えば、外生変数でなければ共分散は仮定できません

外生変数間に共分散がある場合にパラメータが含まれます

構造方程式モデル

- 最後の山ですが、ここまで行列がわからなければ、その意味はよくわかりません。
- 最初に言ったように、測定方程式と構造方程式を合体させたものなので、行列式も両者を合体させたものです。

$$\begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{v} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_a & \boldsymbol{A}_d \\ \boldsymbol{A}_b & \boldsymbol{A}_c \end{bmatrix} \begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{v} \end{bmatrix} + \begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{e} \end{bmatrix}$$

ここで、

$\boldsymbol{d}$: $\boldsymbol{f}$に関する残差変数

$\boldsymbol{e}$: $\boldsymbol{v}$に関する残差変数

$\boldsymbol{A}_a$: f_jからf_iへの係数α_{aij}

$\boldsymbol{A}_b$: f_jからx_iへの係数α_{bij}(因子負荷行列)

$\boldsymbol{A}_c$: x_jからx_iへの係数α_{cij}

$\boldsymbol{A}_d$: x_jからf_iへの係数α_{dij}

構造方程式モデル

- 測定方程式も、構造方程式もこの特殊なケースとなります。

測定方程式

$$\begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{v} \end{bmatrix} = \begin{bmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{A}_b & \boldsymbol{O} \end{bmatrix} \begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{v} \end{bmatrix} + \begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{e} \end{bmatrix}$$

構造方程式

$$\begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{v} \end{bmatrix} = \begin{bmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_c \end{bmatrix} \begin{bmatrix} \boldsymbol{o} \\ \boldsymbol{v} \end{bmatrix} + \begin{bmatrix} \boldsymbol{o} \\ \boldsymbol{e} \end{bmatrix}$$

- 共分散構造

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{O} \end{bmatrix}$$

$$\boldsymbol{T} = (\boldsymbol{I} - \boldsymbol{A})^{-1}$$

$$\boldsymbol{\Sigma} = \boldsymbol{G}\boldsymbol{T}\boldsymbol{\Sigma}_u\boldsymbol{T}'\boldsymbol{G}'$$

$$\boldsymbol{\Sigma}_u = \begin{bmatrix} \boldsymbol{\Sigma}_d & \boldsymbol{\Sigma}_{de} \\ \boldsymbol{\Sigma}_{ed} & \boldsymbol{\Sigma}_e \end{bmatrix}$$

識別問題

- 連立方程式には不能（解が存在しない）と不定（解が無数に存在する）があります。
- 不能の場合は解が存在しませんが、近似解の推定によって母数を求めます。というか、無理やり連立方程式を作っているので、ほとんどこの不能であることは確かです。
- 不定の場合、この方程式は「識別」できません。
- 「十分条件」をクリアすれば方程式は識別できます。十分条件とは、「それが満たされればモデルは識別されるが、満たされないからといってもモデルが識別されるとは限らないという条件」です。
- 一方、「それが満たされればモデルは確実に識別されず、それが満たされるからといってモデルが必ず識別されるとは限らない条件」を必要条件といいます。

SEMのコツ

- Amosを動かしていて悩まされるのがこの識別問題です。
- 教科書によれば、以下の3つが識別を行うコツだそうです。
  1. 十分条件による識別
  2. ソフトウエアによる識別（これは力技です。）
  3. ノウハウによる識別
- 十分条件による識別を行えばモデルは必ず識別されます。（広がりは少ないが・・・）
  - 構成概念 f をいくつか用意
  - 一つの構成概念だけを測定する観測変数をおのおの3つ以上づつ用意
  - 各々の構成概念に関して、それを測定している観測変数から任意に1つ選んでその観測変数への係数を1に固定。構成概念が外生変数なら、分散を1に固定（逆に言えば、外生変数でなければ分散は設定する必要ない）
  - $\boldsymbol{f}$ と $\boldsymbol{d}$ の間に単方向・両方向のパスを引く。
  - 後はもう少し面倒くさい識別条件があるが、これで行えばだいたい大丈夫のようです。

SEMのコツ・その②

- 教科書によれば、こんなノウハウによるコツが紹介されています。
  - すべての残差変数(外生変数 f_i, v_i, 誤差変数 e_i, d_i)には分散を設定
  - 外生的な複数の構造変数の間には共分散を設定(f_iと f_j、v_i とv_j、f_iとv_j)→事前情報に反しない限り
  - 外生的な構造変数には誤差変数が刺さらない
  - 誤差変数間の共分散、誤差変数と外生的な構造変数との間には共分散を設定しない→事前情報に反しない限り
  - 内生的な観測変数には1つ1つ誤差変数がささる→ e_i
  - 内生的な構成概念には1つ1つ誤差変数がささる→ d_i
  - 内生的な変数の分散は設定しない(Amosではもともとできない?)
  - 内生変数間・内生変数と外生変数間の共分散は設定しない
  - モデル中の推定すべき母数の総数(自由度パラメータ)は、観測変数の分散と共分散の和 $n_x(n_x+1)/2$ を超えない
  - f_iの各々に関し、そこからでている単方向の矢を任意に一つ選んでその係数の値を1に固定する。標準化解ならば問題ない。外生変数の時は分散を固定。

適合度指標

- 一時間で終わるために最後ははしょります
  - χ^2検定・・・まああまり役に立たないと割り切ったほうがいいと思います。
  - GFI・・・簡単に言うと、母数によって表現された共分散と、データによる共分散の差です。1が最もよく、0.9以上必要とのことですが、自由度が大きくなると母数が少ない時には数字があがらないとのこと。無理に上げなくてもいいかも。あるいは母数が多ければGFIもあがります。これはR2に似ていますね。教科書は、観測変数の数を少なくしろといっています。（30以下）
  - RMR・・・残差平方平均平方根＝まあ、残差のことですね。0が最もよい。
  - AGFI・・・自由度修正済みのGFI。修正済みR2みたいなもの。
  - CFI・・・比較適合度指標。0から1までの範囲に収まり、1がもっともよい。
  - AIC・・・ご存知、ですが、複数モデルを比較するときに用いるとよい。
- 他にもごちゃごちゃたくさんありますが省略。最初は観測変数を少なくしてGFIだけ使っていれば大丈夫だと思います。また、母数の検定はできますのでこれはAmosを参照してください。