Orthogonal decompositions in the growth curve model

Daniel Klein

Univerzita P.J. Šafárika, Košice

Robust 2010
31. 5.–5. 2. 2010, Králíky

- The idea of orthogonal decompositions in the growth curve model appeared in 2009 in the papers of chinense mathematicians Jianhua Hu, Ren-Dao Ye a Song-Gui Wang. Many tasks, which are very difficult or impossible to handle in basic models, can be done with ease in models consisting of mutually orthogonal components.

- The idea of orthogonal decompositions in the growth curve model appeared in 2009 in the papers of chinense mathematicians Jianhua Hu, Ren-Dao Ye a Song-Gui Wang. Many tasks, which are very difficult or impossible to handle in basic models, can be done with ease in models consisting of mutually orthogonal components.
- Simple transformation can change a model into an equivalent which allows to determine explicit forms of estimators and/or their distribution.

Introduction

- The basic model is the growth curve model introduced in 1964 by Potthoff & Roy. It often appears in many areas (medicine, biology, psychology, economics, ...).

Introduction

- The basic model is the growth curve model introduced in 1964 by Potthoff & Roy. It often appears in many areas (medicine, biology, psychology, economics, ...).
- Model:

$$Y = XBZ + e\,,$$

where

Introduction

- The basic model is the growth curve model introduced in 1964 by Potthoff & Roy. It often appears in many areas (medicine, biology, psychology, economics, ...).
- Model:

$$Y = XBZ + e\,,$$

  where
  - $Y_{n \times p}$ is matrix of p-variate independent observations,

- The basic model is the growth curve model introduced in 1964 by Potthoff & Roy. It often appears in many areas (medicine, biology, psychology, economics, ...).
- Model:

$$Y = XBZ + e\,,$$

  where
  - $Y_{n\times p}$ is matrix of p-variate independent observations,
  - $X_{n\times m}$ and $Z_{r\times p}$ are known design matrices (X is ANOVA matrix and Z is a matrix of regression constants),

Introduction

- The basic model is the growth curve model introduced in 1964 by Potthoff & Roy. It often appears in many areas (medicine, biology, psychology, economics, ...).
- Model:

$$Y = XBZ + e\,,$$

where

  - $Y_{n\times p}$ is matrix of p-variate independent observations,
  - $X_{n\times m}$ and $Z_{r\times p}$ are known design matrices (X is ANOVA matrix and Z is a matrix of regression constants),
  - $B_{m\times r}$ is a matrix of unknown parameters of first order,

- The basic model is the growth curve model introduced in 1964 by Potthoff & Roy. It often appears in many areas (medicine, biology, psychology, economics, ...).
- Model:

$$Y = XBZ + e\,,$$

where

  - $Y_{n\times p}$ is matrix of p-variate independent observations,
  - $X_{n\times m}$ and $Z_{r\times p}$ are known design matrices (X is ANOVA matrix and Z is a matrix of regression constants),
  - $B_{m\times r}$ is a matrix of unknown parameters of first order,
  - $e_{n\times p}$ is an error matrix.

Introduction

- Model:

$$Y = XBZ + e\,,$$

Introduction

- Model:

$$Y = XBZ + e\,,$$

$$\mathrm{E}\, e = 0\,,$$

Introduction

- Model:

$$Y = XBZ + e\,,$$

$$\begin{aligned} \mathrm{E}\, e &= 0\,, \\ \mathrm{Var}(\mathrm{vec}\, e) &= \Sigma \otimes I\,. \end{aligned}$$

vec operator stacks elements of a matrix into a vector column-wise and $\otimes$ denotes the Kroneckerov product of matrices.

- Model:

$$Y = XBZ + e\,,$$

$$\begin{aligned} \mathrm{E}\,e &= 0\,, \\ \mathrm{Var}(\mathrm{vec}\,e) &= \Sigma \otimes I\,. \end{aligned}$$

vec operator stacks elements of a matrix into a vector column-wise and $\otimes$ denotes the Kroneckerov product of matrices.

- We assume that $e'_1, \ldots, e'_n$ are p-variate independent normally distributed vectors with variance matrix Σ. Here e_i denotes the row of matrix e.

Uniform correlation structure

- There is no problem estimating Σ when it is completely unknown. Under normality, its uniformly minimum variance unbiased invariant estimator (UMVUIE) is

$$S = \frac{1}{n - r(X)} Y' M_X Y,$$

where $M_X = I - X(X'X)^{-}X'$.

Uniform correlation structure

- There is no problem estimating Σ when it is completely unknown. Under normality, its uniformly minimum variance unbiased invariant estimator (UMVUIE) is

$$S = \frac{1}{n - r(X)} Y' M_X Y,$$

where $M_X = I - X(X'X)^{-}X'$.

- Problems arise in situations when the structure is partially known. One of the most common structures is the uniform correlation structure:

$$\Sigma = \sigma^2((1-\rho)I_p + \rho \mathbf{1}_p \mathbf{1}_p'),$$

where $\sigma^2 > 0$ and $\rho \in \left\langle -\frac{1}{p-1},\ 1 \right\rangle$ are unknown parameters.

Uniform correlation structure

- Žežula (2006) introduced simple estimators of both parameters based on S:

$$\hat{\sigma}_S^2 = \frac{\mathrm{Tr}(S)}{p},$$
$$\hat{\rho}_S = \frac{1}{p-1}\left(\frac{\mathbf{1}'S\mathbf{1}}{\mathrm{Tr}(S)} - 1\right).$$

Uniform correlation structure

- Žežula (2006) introduced simple estimators of both parameters based on S:

$$\hat{\sigma}_S^2 = \frac{\mathrm{Tr}(S)}{p},$$
$$\hat{\rho}_S = \frac{1}{p-1}\left(\frac{\mathbf{1}'S\mathbf{1}}{\mathrm{Tr}(S)} - 1\right).$$

- Both estimators are based on unbiased estimating equations, however, the estimator $\hat{\rho}_S$ is biased, and the boundaries are $-\frac{1}{p-1} \leq \hat{\rho}_S \leq 1$

Uniform correlation structure

- The properties of estimator of $\hat{\rho}_S$:

Uniform correlation structure

- The properties of estimator of $\hat{\rho}_S$:
  - $\mathrm{E}\,\hat{\rho}_S = \rho - \frac{2}{n-r(X)}\rho(1-\rho)\frac{1+(p-1)\rho}{p} + O(n^{-2})$,

Uniform correlation structure

- The properties of estimator of $\hat{\rho}_S$:
  - $\mathrm{E}\,\hat{\rho}_S = \rho - \frac{2}{n-r(X)}\rho(1-\rho)\frac{1+(p-1)\rho}{p} + O(n^{-2})$,
  - $\mathrm{MSE}\hat{\rho}_S = \frac{2}{n-r(X)} \cdot \frac{(1-\rho)^2[1+(p-1)\rho]^2}{p(p-1)} + O(n^{-2})$,

Uniform correlation structure

- The properties of estimator of $\hat{\rho}_S$:
  - $\mathrm{E}\,\hat{\rho}_S = \rho - \frac{2}{n-r(X)}\rho(1-\rho)\frac{1+(p-1)\rho}{p} + O(n^{-2})$,
  - $\mathrm{MSE}\hat{\rho}_S = \frac{2}{n-r(X)} \cdot \frac{(1-\rho)^2[1+(p-1)\rho]^2}{p(p-1)} + O(n^{-2})$,
  - the distribution of $\hat{\rho}_S$ was an open problem for a long time

Uniform correlation structure

- The properties of estimator of $\hat{\rho}_S$:
  - $\mathrm{E}\,\hat{\rho}_S = \rho - \frac{2}{n-r(X)}\rho(1-\rho)\frac{1+(p-1)\rho}{p} + O(n^{-2})$,
  - $\mathrm{MSE}\hat{\rho}_S = \frac{2}{n-r(X)} \cdot \frac{(1-\rho)^2[1+(p-1)\rho]^2}{p(p-1)} + O(n^{-2})$,
  - the distribution of $\hat{\rho}_S$ was an open problem for a long time
- Using the transformation

$$Z_n = \frac{1}{2}\ln\left(\frac{\frac{1}{p-1}+\hat{\rho}_S}{1-\hat{\rho}_S}\right),$$

asymptotic normality can be achieved.

- Žežula $\longrightarrow$ used directly the basic model.
  Ye&Wang $\longrightarrow$ an idea to use not directly basic model but modified model with orthogonal decomposition:

  $$Y = Y_1 + Y_2,$$

  where

  $$Y_1 = YP_{\mathbf{1}} = XBZ'P_{\mathbf{1}} + e_1,$$
  $$Y_2 = YM_{\mathbf{1}} = XBZ'M_{\mathbf{1}} + e_2.$$

  Here $\mathbf{1} = (1, \ldots, 1)'$ a $P_{\mathbf{1}} = \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'$, $M_{\mathbf{1}} = I - P_{\mathbf{1}}$.

Uniform correlation structure

- Then the estimators are based on

$$V_1 = P_{\mathbf{1}} S P_{\mathbf{1}}, \quad V_2 = M_{\mathbf{1}} S M_{\mathbf{1}}\,,$$

$S = \frac{1}{n-r(X)} Y' M_X Y$ is UMVUIE of Σ, the estimator on which are based $\hat{\sigma}^2_S$ a $\hat{\rho}_S$.

Uniform correlation structure

- Then the estimators are based on

$$V_1 = P_{\mathbf{1}} S P_{\mathbf{1}}, \quad V_2 = M_{\mathbf{1}} S M_{\mathbf{1}} \, ,$$

$S = \frac{1}{n-r(X)} Y' M_X Y$ is UMVUIE of Σ, the estimator on which are based $\hat{\sigma}^2_S$ a $\hat{\rho}_S$.

- The formulas:

$$\hat{\sigma}^2_{YW} = \frac{\operatorname{Tr}(V_1) + \operatorname{Tr}(V_2)}{p} \, ,$$

$$\hat{\rho}_{YW} = 1 - \frac{p \operatorname{Tr}(V_2)}{(p-1)(\operatorname{Tr}(V_1) + \operatorname{Tr}(V_2))} \, ,$$

Uniform correlation structure

- It is easy to show that this estimators are equivalent, i.e. $\hat{\sigma}^2_S = \hat{\sigma}^2_{YW}$ and $\hat{\rho}_S = \hat{\rho}_{YW}$ for any Y.

Uniform correlation structure

- It is easy to show that this estimators are equivalent, i.e. $\hat{\sigma}^2_S = \hat{\sigma}^2_{YW}$ and $\hat{\rho}_S = \hat{\rho}_{YW}$ for any Y.
- However, orthogonal decomposition is very useful for derivation of the distribution of estimators in an easy way:
  Distributions of $\operatorname{Tr}(V_1)$ and $\operatorname{Tr}(V_2)$ are independent,

$$\operatorname{Tr}(V_1) \sim \frac{\sigma^2[1+(p-1)\rho]}{n-r(X)}\,\chi^2_{n-r(X)},$$

$$\operatorname{Tr}(V_2) \sim \frac{\sigma^2(1-\rho)}{n-r(X)}\,\chi^2_{(p-1)(n-r(X))},$$

  so that

$$\hat{\sigma}^2 \sim \frac{\sigma^2}{p(n-r(X))}\left[(1+(p-1)\rho)\chi^2_{n-r(X)} + (1-\rho)\chi^2_{(p-1)(n-r(X))}\right],$$

$$\frac{1-\rho}{1+(p-1)\rho}\left[\frac{1+(p-1)\hat{\rho}}{1-\hat{\rho}}\right] \sim F_{n-r(X),(p-1)(n-r(X))}\,.$$

- This result is not very useful with respect to $\hat{\sigma}^2$, its distribution depends on both σ^2 and ρ, but enables us to test for any specific value of ρ.

  $100(1-\alpha)\%$ confidence interval for ρ is given

$$\left(\frac{1-c_1}{1+(p-1)c_1};\frac{1-c_2}{1+(p-1)c_2}\right),$$

  where

$$c_1=\frac{1-\hat{\rho}}{1+(p-1)\hat{\rho}}\,F_{n-r(X),(p-1)(n-r(X))}\left(1-\frac{\alpha}{2}\right)$$

  and

$$c_2=\frac{1-\hat{\rho}}{1+(p-1)\hat{\rho}}\,F_{n-r(X),(p-1)(n-r(X))}\left(\frac{\alpha}{2}\right).$$

Example

- Consider random sample from bivariate normal distribution with the same variances in both dimensions.

Example

- Consider random sample from bivariate normal distribution with the same variances in both dimensions.
- Formally it can be written as GCM model with the uniform correlation structure:

$$Y = \begin{pmatrix} Y_{11} & Y_{12} \\ \vdots & \vdots \\ Y_{n1} & Y_{n2} \end{pmatrix} = \mathbf{1}_n \left(\mu_1, \mu_2\right) I_2 + e \,,$$

$$e \sim N_{n\times 2}\left(0_{n\times 2}, \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \otimes I_n\right) .$$

Example

- Using the above mentioned estimator we get

$$\hat{\rho} = \frac{2s_{12}}{s_1^2 + s_2^2},$$

where s_{12} is sample covariance of the two variables, and s_1^2 and s_2^2 sample variances.

Example

- Using the above mentioned estimator we get

$$\hat{\rho} = \frac{2s_{12}}{s_1^2 + s_2^2},$$

  where s_{12} is sample covariance of the two variables, and s_1^2 and s_2^2 sample variances.
- This estimator is slightly more effective than standard sample correlation coefficient (in the sense of MSE).

Linearly structured covariance matrix

- In fact, also M. Ohlson used orthogonal decomposition in deriving the explicit estimators of linearly structured covariance matrix.

- In fact, also M. Ohlson used orthogonal decomposition in deriving the explicit estimators of linearly structured covariance matrix.
- $\Sigma = (\sigma_{ij})$ is linearly structured if the only linear structure between the elements is given by $|\sigma_{ij}| = |\sigma_{kl}|$ and there exists at least one $(i, j) \neq (k, l)$ so that $|\sigma_{ij}| = |\sigma_{kl}|$.
  EXAMPLE: Toeplitz structure

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho_1 & \rho_2 \\ \rho_1 & \sigma^2 & \rho_1 \\ \rho_2 & \rho_1 & \sigma^2 \end{pmatrix}$$

Linearly structured covariance matrix

- In fact, also M. Ohlson used orthogonal decomposition in deriving the explicit estimators of linearly structured covariance matrix.
- $\Sigma = (\sigma_{ij})$ is linearly structured if the only linear structure between the elements is given by $|\sigma_{ij}| = |\sigma_{kl}|$ and there exists at least one $(i,j) \neq (k,l)$ so that $|\sigma_{ij}| = |\sigma_{kl}|$.
  EXAMPLE: Toeplitz structure

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho_1 & \rho_2 \\ \rho_1 & \sigma^2 & \rho_1 \\ \rho_2 & \rho_1 & \sigma^2 \end{pmatrix}$$

- MLEs in such model must be found numerically.

Linearly structured covariance matrix

- Ohlson used the decomposition similar to Ye&Wang, but he made use the decomposition of the space $\mathcal{M}(X)$:

$$Y = Y_1 + Y_2,$$

where

$$Y_1 = P_X Y = XBZ + e_1,$$
$$Y_2 = M_X Y = e_2.$$

Linearly structured covariance matrix

- Ohlson used the decomposition similar to Ye&Wang, but he made use the decomposition of the space $\mathcal{M}(X)$:

$$Y = Y_1 + Y_2,$$

where

$$Y_1 = P_X Y = XBZ + e_1,$$
$$Y_2 = M_X Y = e_2.$$

- The idea is to decompose the total variation into two terms: $(n - r(X))S$ and $\hat{R}_1'\hat{R}_1$, where

$$\hat{R}_1 = P_X Y(M_{Z'}^{S^{-1}})'.$$

Linearly structured covariance matrix

- Ohlson used the decomposition similar to Ye&Wang, but he made use the decomposition of the space $\mathcal{M}(X)$:

$$Y = Y_1 + Y_2,$$

where

$$Y_1 = P_X Y = XBZ + e_1,$$
$$Y_2 = M_X Y = e_2.$$

- The idea is to decompose the total variation into two terms: $(n - r(X))S$ and $\hat{R}_1' \hat{R}_1$, where

$$\hat{R}_1 = P_X Y (M_{Z'}^{S^{-1}})'.$$

- The total variation is the sums of this two terms.

Linearly structured covariance matrix

- The first step: using the variation as when obtaining the estimator for Σ in unstructured case, in fact, it is the estimator of Σ in the model Y_2 using least square method:

$$\operatorname{vec}\hat{\Sigma}_1 = T'(TT')^{-}T \operatorname{vec} S,$$

where T is matrix so that $T \operatorname{vec}\Sigma = \operatorname{vec}\Sigma(K)$. Here $\operatorname{vec}\Sigma(K)$ is the columnwise vectorized form of Σ where all 0 and repeated elements have been disregarded.

Linearly structured covariance matrix

- The first step: using the variation as when obtaining the estimator for Σ in unstructured case, in fact, it is the estimator of Σ in the model Y_2 using least square method:

$$\operatorname{vec}\hat{\Sigma}_1 = T'(TT')^{-}T\operatorname{vec}S,$$

where T is matrix so that $T\operatorname{vec}\Sigma = \operatorname{vec}\Sigma(K)$. Here $\operatorname{vec}\Sigma(K)$ is the columnwise vectorized form of Σ where all 0 and repeated elements have been disregarded.

- The final estimator is derived by the least square method using the total variation $(n - r(X))S + \hat{R}_1'\hat{R}_1$, with

$$\hat{R}_1 = P_X Y(M_{Z'}^{\hat{\Sigma}_1^{-1}})'.$$

Extended model

- The extended growth curve model (ECGM) with fixed effects (sum-of-profiles model) is

$$Y = \sum_{i=1}^{k} X_i B_i Z_i' + e\,, \quad e \sim N_{n\times p}\left(0, \Sigma \otimes I_n\right).$$

Extended model

- The extended growth curve model (ECGM) with fixed effects (sum-of-profiles model) is

$$Y = \sum_{i=1}^{k} X_i B_i Z_i' + e \,, \quad e \sim N_{n \times p}\left(0, \Sigma \otimes I_n\right).$$

- Usually it is supposed that the column spaces of X_i's are ordered

$$\mathcal{M}(X_k) \subseteq \cdots \subseteq \mathcal{M}(X_1),$$

and nothing is said about different Z_i's.

Extended model

- The extended growth curve model (ECGM) with fixed effects (sum-of-profiles model) is

$$Y = \sum_{i=1}^{k} X_i B_i Z_i' + e\,, \quad e \sim N_{n\times p}\left(0, \Sigma \otimes I_n\right).$$

- Usually it is supposed that the column spaces of X_i's are ordered

$$\mathcal{M}(X_k) \subseteq \cdots \subseteq \mathcal{M}(X_1),$$

and nothing is said about different Z_i's.

- Now, the idea is to separate groups rather than models and so consider

$$X_i' X_j = 0 \;\; \forall i \neq j.$$

Extended model

- Since we consider that X_i's are ANOVA matrices and all X_i's and Z_i's are of full rank, it can be shown that the two models are equivalent:

- Since we consider that X_i's are ANOVA matrices and all X_i's and Z_i's are of full rank, it can be shown that the two models are equivalent:
  - X_i's are $0-1$ matrices whose columns are indicators of different groups,

Extended model

- Since we consider that X_i's are ANOVA matrices and all X_i's and Z_i's are of full rank, it can be shown that the two models are equivalent:
  - X_i's are $0-1$ matrices whose columns are indicators of different groups,
  - we can assume that all columns of X_1 are perpendicular and columns of every X_i are a subset of columns of X_{i-1},

Extended model

- Since we consider that X_i's are ANOVA matrices and all X_i's and Z_i's are of full rank, it can be shown that the two models are equivalent:
  - X_i's are $0-1$ matrices whose columns are indicators of different groups,
  - we can assume that all columns of X_1 are perpendicular and columns of every X_i are a subset of columns of X_{i-1},
  - we define

$$X_k^* = X_k,$$

$$X_i^* = X_i \diagdown X_{i+1}, \quad i = 1, \dots, k-1.$$

Here $X_i \diagdown X_{i+1}$ consists of those columns of X_i which are not in X_{i+1}.

Extended model

- Then we can reformulate the model

$$\sum_{i=1}^{k} X_i B_i Z_i' = \sum_{j=1}^{k} X_j^* B_j^* Z_j^{*\prime},$$

where

$$Z_i^* = (Z_1, \ldots, Z_i), \quad \forall\, i = 1, \ldots, k,$$

which implies $\mathcal{R}(Z_1^*) \subset \cdots \subset \mathcal{R}(Z_k^*)$.

Extended model

EXAMPLE

- Consider EGCM with two groups with different growth patterns – linear and quadratic:

$$
\begin{aligned}
Y_{ij} &= \beta_1 + \beta_2 t_j + e_{ij}\,, && i = 1, \ldots, n_1\,,\ j = 1, \ldots, p\,,\\
&= \beta_3 + \beta_4 t_j + \beta_5 t_j^2 + e_{ij}\,, && i = n_1 + 1, \ldots, n_1 + n_2\,,\ j = 1, \ldots, p\,.
\end{aligned}
$$

Extended model

EXAMPLE

- Consider EGCM with two groups with different growth patterns – linear and quadratic:

$$\begin{aligned} Y_{ij} &= \beta_1 + \beta_2 t_j + e_{ij}\,, && i = 1, \dots, n_1\,,\ j = 1, \dots, p\,, \\ &= \beta_3 + \beta_4 t_j + \beta_5 t_j^2 + e_{ij}\,, && i = n_1 + 1, \dots, n_1 + n_2\,,\ j = 1, \dots, p\,. \end{aligned}$$

- This model can be written as

$$Y = \begin{pmatrix} \mathbf{1}_{n_1} & 0 \\ 0 & \mathbf{1}_{n_2} \end{pmatrix} \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} 1 & \dots & 1 \\ t_1 & \dots & t_p \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{1}_{n_2} \end{pmatrix} \beta_5 \begin{pmatrix} t_1^2 & \dots & t_p^2 \end{pmatrix} + e,$$

or, by the new way, as

$$Y = \begin{pmatrix} \mathbf{1}_{n_1} \\ 0 \end{pmatrix} (\beta_1, \beta_2) \begin{pmatrix} 1 & \dots & 1 \\ t_1 & \dots & t_p \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{1}_{n_2} \end{pmatrix} (\beta_3, \beta_4, \beta_5) \begin{pmatrix} 1 & \dots & 1 \\ t_1 & \dots & t_p \\ t_1^2 & \dots & t_p^2 \end{pmatrix} + e.$$

Extended model

- The model with a condition of perpendicular X_i's is much easier to handle.

Extended model

- The model with a condition of perpendicular X_i's is much easier to handle.
- If all X_i^*'s and Z_i^*'s are of full rank, then all B_i^*'s are estimable and unbiased LSE

$$\hat{B}_i^* = \left(X_i^{*\prime} X_i^*\right)^{-1} X_i^{*\prime} Y \Sigma^{-1} Z_i^* \left(Z_i^{*\prime} \Sigma^{-1} Z_i^*\right)^{-1}$$

depend only on X_i^* and Z_i^*.

Extended model

- Such a closed form was difficult to obtain in the model with ordered spaces of X_i's. Even for two components the estimators are rather complicated:

$$\hat{B}_1 = \left(X_1{}'X_1\right)^{-1} X_1{}'Y\Sigma^{-1}Z_1\left(Z_1{}'\Sigma^{-1}Z_1\right)^{-1} - \left(X_1{}'X_1\right)^{-1} X_1{}'P_{X_2}Y\left(P_{Z_2}^{\Sigma^{-1}M_{Z_1}^{\Sigma^{-1}}}\right)'\Sigma^{-1}Z_1\left(Z_1{}'\Sigma^{-1}Z_1\right)^{-1},$$

$$\hat{B}_2 = \left(X_2{}'X_2\right)^{-1} X_2{}'Y\Sigma^{-1}Z_2\left(Z_2{}'\Sigma^{-1}M_{Z_1}^{\Sigma^{-1}}Z_2\right)^{-1},$$

Extended model

- Such a closed form was difficult to obtain in the model with ordered spaces of X_i's. Even for two components the estimators are rather complicated:

$$\hat{B}_1 = \left(X_1'X_1\right)^{-1} X_1'Y\Sigma^{-1}Z_1 \left(Z_1'\Sigma^{-1}Z_1\right)^{-1} - \left(X_1'X_1\right)^{-1} X_1'P_{X_2}Y \left(P_{Z_2}^{\Sigma^{-1}M_{Z_1}^{\Sigma^{-1}}}\right)' \Sigma^{-1}Z_1 \left(Z_1'\Sigma^{-1}Z_1\right)^{-1},$$

$$\hat{B}_2 = \left(X_2'X_2\right)^{-1} X_2'Y\Sigma^{-1}Z_2 \left(Z_2'\Sigma^{-1}M_{Z_1}^{\Sigma^{-1}}Z_2\right)^{-1},$$

- $\hat{B}_1$ and $\hat{B}_2$ depends on both Z_1 and Z_2, and $\hat{B}_1$ even on X_2.

Extended model

- Also using Ohlson method the explicit estimator of linearly structured covariance matrix can be found in an easy way:

$$\operatorname{vec}\hat{\Sigma} = T^{+}\left(\left(T^{+}\right)'\hat{\Psi}'\hat{\Psi}T^{+}\right)'\hat{\Psi}'\operatorname{vec}(Y-\hat{Y})'(Y-\hat{Y})$$

  where

$$\hat{Y} = P_{(X_1,\dots,X_k)}Y\left(P_{Z_1'}^{S^{-1}} + \cdots + P_{Z_k'}^{S^{-1}}\right)',$$

$$\hat{\Psi} = \sum_{i=1}^{k} r(X_i)\,M_{Z_i'}^{S^{-1}} \otimes M_{Z_i'}^{S^{-1}} + (n - \sum r(X_i))I.$$

Thank you for your attention