Truncated partial correlation vine approach for recursive structural equation models

Outline:

- Factor-longitudinal example in Mulaik's book: simplex structure and factor-tree structure
- Expressing the models in the `lavaan` R library (see Rosseel (2012), *J Stat Software* 48(2), for goals of `lavaan` and comparisons to proprietary software such as EQS, AMOS, Mplus.
- Estimation with truncated partial correlation vine

Example for simplex structure

values decreasing along the diagonals (inner to outer) = Simplex. Guttman's idea.

Correlation matrix of eight stages of the "Practice in the Discrimination Reaction Time" test, $n = 264$; page 271 of Mulaik's book. Note the decrease in correlation with lag. The last column has the row sum

1-Factor dependence structure
[1 a1a2 a1a3 … a1aD
a2a1 1 a2a3 ….. a2aD]

where a1a2 greater than a1a3

If there is major deviation from 1 fd structure, in the simplex model, can add more edges to your fd structure.

1.00	0.74	0.71	0.69	0.62	0.59	0.57	0.56	5.48
0.74	1.00	0.82	0.78	0.74	0.68	0.66	0.64	6.06
0.71	0.82	1.00	0.83	0.79	0.72	0.72	0.71	6.30
0.69	0.78	0.83	1.00	0.80	0.74	0.72	0.74	6.30
0.62	0.74	0.79	0.80	1.00	0.77	0.73	0.77	6.22
0.59	0.68	0.72	0.74	0.77	1.00	0.74	0.80	6.04
0.57	0.66	0.72	0.72	0.73	0.74	1.00	0.79	5.93
0.56	0.64	0.71	0.74	0.77	0.80	0.79	1.00	6.01

What dependence structures can be considered?

Longitudinal dependence comes from time-ordered variables (correlation matrix is decreasing along diagonals away from main diagonal but not geometrically decreasing). ? Number at 1,3 will be the product of (1,2) and (2,3)

Factor structure come from repeated measures.

By inspection with similarly ordered rows ($\rho_{jm} - \rho_{km}$ mostly have same sign for $m \neq j, k$), a 1-factor fit suggests that the loading for item3 is larger than that of item2 which in turn is larger than that of item1. Also row 4 partly dominates rows 6,7,8.

An example of an output from Lavaan:

1.00	0.74	0.71	0.69	0.62	0.59	0.57	0.56	0.000	0.104	0.039	0.017	-0.041	-0.040	-0.045	-0.066
0.74	1.00	0.82	0.78	0.74	0.68	0.66	0.64	0.104	0.000	0.054	0.012	-0.014	-0.038	-0.041	-0.074
0.71	0.82	1.00	0.83	0.79	0.72	0.72	0.71	0.039	0.054	0.000	0.019	-0.007	-0.039	-0.021	-0.045
0.69	0.78	0.83	1.00	0.80	0.74	0.72	0.74	0.017	0.012	0.019	0.000	0.001	-0.020	-0.023	-0.016
0.62	0.74	0.79	0.80	1.00	0.77	0.73	0.77	-0.041	-0.014	-0.007	0.001	0.000	0.022	0.000	0.026
0.59	0.68	0.72	0.74	0.77	1.00	0.74	0.80	-0.040	-0.038	-0.039	-0.020	0.022	0.000	0.045	0.092
0.57	0.66	0.72	0.72	0.73	0.74	1.00	0.79	-0.045	-0.041	-0.021	-0.023	0.000	0.045	0.000	0.099
0.56	0.64	0.71	0.74	0.77	0.80	0.79	1.00	-0.066	-0.074	-0.045	-0.016	0.026	0.092	0.099	0.000

1-factor fit (below), $\boldsymbol{R}_{\text{obs}}$ (above left), and residual difference $\boldsymbol{R}_{\text{obs}} - \boldsymbol{R}_{\text{1factor}}$ (above, right)

1.000	0.636	0.671	0.673	0.661	0.630	0.615	0.626
0.636	1.000	0.766	0.768	0.755	0.718	0.702	0.715
0.671	0.766	1.000	0.811	0.797	0.759	0.741	0.755
0.673	0.768	0.811	1.000	0.799	0.760	0.743	0.756
0.661	0.755	0.797	0.799	1.000	0.748	0.730	0.743
0.630	0.718	0.759	0.760	0.748	1.000	0.695	0.708
0.615	0.702	0.741	0.743	0.730	0.695	1.000	0.691
0.626	0.715	0.755	0.756	0.743	0.708	0.691	1.000

Factor tree model says there is a small amount of conditional dependence in Y1,…Yd | W.

1-factor fit and residual matrix

W
|||||||
Y1 <Py1y2, Py1y5> Y2 <Py2y3, Py2Y8> Y3 <> Y4 <> Y5 <> Y6 <> Y7 <> Y8

ordered (decreasing) absolute "residuals" from previous slide: 0.104 0.099 0.092 0.074 0.066 0.054 0.045 0.045 0.045 0.041 0.041 0.040 0.039 0.039 0.038 0.026 ... 0.023

Corresponding pairs are: 12, 78, 68, 28, [18], 23, [67, 17, 27], 15, ... 47

A tree can be formed for the conditional dependence of variables given the latent variable.

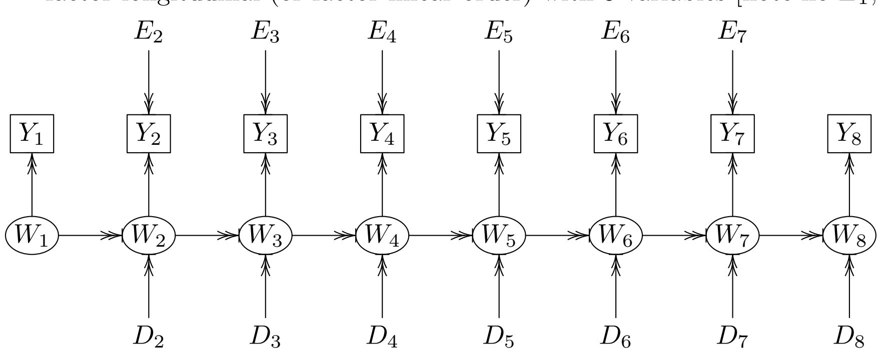
This would correspond to a factor-tree model as a 2-truncated vine (forthcoming slide).

Graphs of models (include the disturbances in one case)

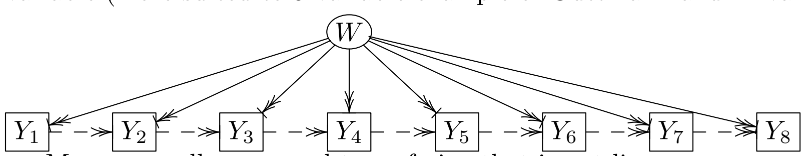
Simplex structure for longitudinal or for variables that can be linearly ordered (latent AR). $Y_1, \dots, Y_d$ are ordered by time or difficulty, etc. W_j is latent variable associated to Y_j. $\{W_1, \dots, W_d\}$ is AD(1) or Markov order 1.

(see psychcor.pdf for Guttman's example where $d = 9$ test scores Y_j are ordered by degree of difficulty)

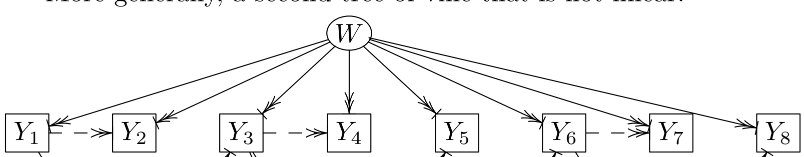
factor-longitudinal (or factor-linear-order) with 8 variables [note no E_1, E_8]



Alternatively, one latent factor and one layer of conditional dependence of observed variables given the latent variable (more suited to 9-variable example of Guttmann and 21-variable example of Thurstone[2]).



More generally, a second tree of vine that is not linear.



Summary of discrepancy of fit, max and average absolute deviation observed and model-fitted correlations

model	Dfit	maxabs	avgabs	#par	
1factor	0.546	0.104	0.038	8	
2factor	0.046	0.027	0.007	15	
simplex	0.047	0.038	0.008	13	***
1factor-linear	0.152	0.066	0.017	15	Linear is not the best way to handle the residual.
1factor-tree1	0.061	0.031	0.011	15	
1factor-tree2	0.045	0.032	0.009	15	
1truncated	0.938	0.345	0.163	7	
2truncated	0.220	0.111	0.038	13	

2-factor / varimax: each factor loads on each observed variable.

tree1 from residual matrix

tree2 from optimal tree, Brechmann-Joe (2014) lavaan notation: =∼ for "is manifested in", ∼ for linear regression relation, ∼∼ for (residual) variance or covariance (disturbance variable).

Simplex in lavaan ('u' for unobserved)

```
# latent variables
    uv1 =~ v1; uv2 =~ v2; uv3 =~ v3; uv4 =~ v4;
    uv5 =~ v5; uv6 =~ v6; uv7 =~ v7; uv8 =~ v8
# regressions
    uv2 ~ uv1; uv3 ~ uv2; uv4 ~ uv3; uv5 ~ uv4;
    uv6 ~ uv5; uv7 ~ uv6; uv8 ~ uv7
# variances
    v2 ~~ v2; v3 ~~ v3; v4 ~~ v4;
    v5 ~~ v5; v6 ~~ v6; v7 ~~ v7;
    uv2 ~~ uv2; uv3 ~~ uv3; uv4 ~~ uv4;
    uv5 ~~ uv5; uv6 ~~ uv6; uv7 ~~ uv7; uv8 ~~ uv8
```

Factor-lineartree in lavaan

```
# latent variables
    uv =~ v1 + v2 + v3 + v4 + v5 + v6 + v7 + v8
# regressions
    v2 ~ v1 + uv; v3 ~ v2 + uv; v4 ~ v3 + uv; v5 ~ v4 + uv;
    v6 ~ v5 + uv; v7 ~ v6 + uv; v8 ~ v7 + uv
# variances
    v1 ~~ v1; v2 ~~ v2; v3 ~~ v3; v4 ~~ v4;
    v5 ~~ v5; v6 ~~ v6; v7 ~~ v7; v8 ~~ v8;
```

Simplex structure is a Markov tree in 16 variables (8 latent).

Stochastic representation and get the correlation matrix.

$Y_j = \alpha_j W_j + \psi_j \epsilon_j$, $j = 1, \ldots, 8$

$W_j = \phi_j W_{j-1} + \psi_j^* \epsilon_j^*$, $j = 2, \ldots, 8$

Correlation matrix to be derived in class. It is a "product" of factor structure and AR(1) structure.

$$\begin{pmatrix} 1 & \alpha_1\alpha_2\phi_2 & \alpha_1\alpha_3\phi_2\phi_3 & \alpha_1\alpha_4\phi_2\phi_3\phi_4 & \alpha_1\alpha_5\phi_2\cdots\phi_5 & \alpha_1\alpha_6\phi_2\cdots\phi_6 & \alpha_1\alpha_7\phi_2\cdots\phi_7 & \alpha_1\alpha_8\phi_2\cdots\phi_8 \\ & 1 & \alpha_2\alpha_3\phi_3 & \alpha_2\alpha_4\phi_3\phi_4 & \alpha_2\alpha_5\phi_3\cdots\phi_5 & \alpha_2\alpha_6\phi_3\cdots\phi_6 & \alpha_2\alpha_7\phi_3\cdots\phi_7 & \alpha_2\alpha_8\phi_3\cdots\phi_8 \\ & & 3 & \alpha_3\alpha_4\phi_4 & \alpha_3\alpha_5\phi_4\phi_5 & \alpha_3\alpha_6\phi_4\cdots\phi_6 & \alpha_3\alpha_7\phi_4\cdots\phi_7 & \alpha_3\alpha_8\phi_4\cdots\phi_8 \\ & & & 4 & \alpha_4\alpha_5\phi_5 & \alpha_4\alpha_6\phi_5\phi_6 & \alpha_4\alpha_7\phi_5\cdots\phi_7 & \alpha_4\alpha_8\phi_5\cdots\phi_8 \\ & & & & 5 & \alpha_5\alpha_6\phi_6 & \alpha_5\alpha_7\phi_6\phi_7 & \alpha_5\alpha_8\phi_6\cdots\phi_8 \\ & & & & & 1 & \alpha_6\alpha_7\phi_7 & \alpha_6\alpha_8\phi_7\phi_8 \\ & & & & & & 1 & \alpha_7\alpha_8\phi_8 \\ & & & & & & & 1 \end{pmatrix}$$

$\alpha_8\phi_8$ appear together (**easy to see from path diagram**), so take $\alpha_8 = 1$; $\alpha_1\phi_2$ appear together, so take $\alpha_1 = 1$; #parameters=13 (no extra residual for Y_1, Y_8, for example).

Parameter estimates (maximum likelihood) for simplex model (`lavaan` and vine approach)

parameter	estimate	parameter	estimate
$\alpha_1\phi_2$	0.801		
α_2	0.923		
α_3	0.936	ϕ_3	0.949
α_4	0.923	ϕ_4	0.964
α_5	0.913	ϕ_5	0.953
α_6	0.889	ϕ_6	0.947
α_7	0.866	ϕ_7	0.977
$\alpha_8\phi_8$	0.912		

Vine array representation for 1-factor with linear tree

$$A = \begin{bmatrix} W & W & W & W & W & W & W & W & W \\ & Y_1 & Y_1 & Y_2 & Y_3 & Y_4 & Y_5 & Y_6 & Y_7 \\ & & Y_2 & & & & & & \\ & & & Y_3 & & & & & \\ & & & & Y_4 & & & & \\ & & & & & Y_5 & & & \\ & & & & & & Y_6 & & \\ & & & & & & & Y_7 & \\ & & & & & & & & Y_8 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ & 2 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ & & 3 & & & & & & \\ & & & 4 & & & & & \\ & & & & 5 & & & & \\ & & & & & 6 & & & \\ & & & & & & 7 & & \\ & & & & & & & 8 & \\ & & & & & & & & 9 \end{bmatrix}$$

leading variables are the observed variables; ones that follow are the latent ones.

Vine approach:
1. Parametrize in terms of 8 correlations with latent variable and 7 partial correlations of observed given latent.
2. `pcor2cor(pcobj,A,ntrunc=2)` to get correlation matrix of $(W, Y_1, \ldots, Y_8)$. Marginalize to get correlation matrix of $(Y_1, \ldots, Y_8)$. Put into multivariate normal log-likelihood and optimize (`nlm` or `optim`) over the 15 algebraically independent parameters in $(-1, 1)$.
3. At the end, `pcor2cor(pcmle,A,ntrunc=2)` to get correlation matrix and matrix of regression coefficients.

`lavaan` outputs regression coefficients and the "residual matrix" of observed correlation matrix minus model-fitted correlation matrix, and these can be compared.

Estimates of correlations, partial correlations and regression coefficients for factor-lineartree structure (Top table from vine approach, bottom table from vine and `lavaan`).

variable	ρ_{jW}	$\rho_{j,j-1;W}$
Y_1	0.691	.
Y_2	0.793	0.435
Y_3	0.846	0.458
Y_4	0.864	0.369
Y_5	0.900	0.102
Y_6	0.880	-0.109
Y_7	0.825	0.053
Y_8	0.865	0.271

variable	β_{jW}	$\beta_{j,j-1}$
Y_1	0.692	0.000
Y_2	0.540	0.367
Y_3	0.528	0.401
Y_4	0.568	0.349
Y_5	0.824	0.088
Y_6	0.987	-0.119
Y_7	0.770	0.063
Y_8	0.667	0.240

Bi-factor (a common factor for all variables, G group-based factors that each load on a group). #pars=$2d$ (if group sizes > 2).

Tri-factor (a common factor for all variables, G group-based factors that each load on one group, H subgroup-based factors that each load on one subgroup; each group are partitioned into subgroups). #parameters=$3d$ if ... Items are divided into $G \geq 2$ groups. For $g = 1, \ldots, G$, group g is divided into s_g subgroups. The number of items into each subgroup is at least 2.

Bi-factor loading matrix $\boldsymbol{A}$ can be written as

$$\begin{pmatrix} \boldsymbol{\phi_1} & \boldsymbol{\alpha_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \boldsymbol{\phi_2} & \mathbf{0} & \boldsymbol{\alpha_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\phi_G} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\alpha_G} \end{pmatrix},$$

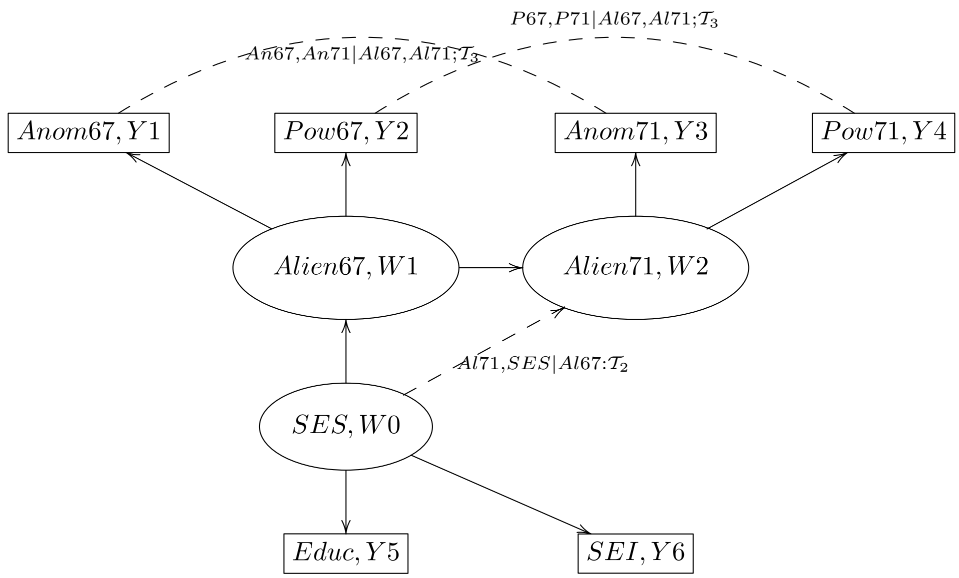
Tri-factor example: a loading matrix for $G = 2$, $s_1 = 3$ subgroups, $s_2 = 4$ subgroups is shown below. The first column has the common latent variable, followed by latent variables that only load on specific groups, followed by latent variables that only load on specific subgroups:

$$\begin{pmatrix} \boldsymbol{\alpha}_{10} & \boldsymbol{\alpha}_{11} & & \boldsymbol{\alpha}_{13} & & & & & & \\ \boldsymbol{\alpha}_{20} & \boldsymbol{\alpha}_{21} & & & \boldsymbol{\alpha}_{24} & & & & & \\ \boldsymbol{\alpha}_{30} & \boldsymbol{\alpha}_{31} & & & & \boldsymbol{\alpha}_{35} & & & & \\ \boldsymbol{\alpha}_{40} & & \boldsymbol{\alpha}_{42} & & & & \boldsymbol{\alpha}_{46} & & & \\ \boldsymbol{\alpha}_{50} & & \boldsymbol{\alpha}_{52} & & & & & \boldsymbol{\alpha}_{57} & & \\ \boldsymbol{\alpha}_{60} & & \boldsymbol{\alpha}_{62} & & & & & & \boldsymbol{\alpha}_{68} & \\ \boldsymbol{\alpha}_{70} & & \boldsymbol{\alpha}_{72} & & & & & & & \boldsymbol{\alpha}_{79} \end{pmatrix}$$

`CopulaModel` has these models implemented efficiently (multivariate Gaussian and t): `factanal.co, factanal.bi, factanal.tri`. The latter two would be tedious(?) to use in lavaan for large d? If there are no subgroups, the tri-factor loading matrix matches the bi-factor model. Because the number of variables can be quite large for these models, the optimization uses analytical derivatives and function evaluations in Fortran90.

Wheaton example in lavaan documentation and Section 7.11 of *Dependence Modeling with Copulas* (code in dmwc-examples-20140627 archive file at copula.stat.ubc.ca)

Attitude scales measuring Anomia and Powerlessness were administered to a sample of 932 people in 1967 and 1971; Y_1 =anomia1967, Y_2 =powerlessness1967, Y_3 =anomia1971, Y_4 =powerlessness1971, Y_5 =education, Y_6 =SEI (socioeconomic index).
Latent variables W_1 =alienation1967 and W_2 =alienation1971, W_0 =SES (socioeconomic status).



To do in class:
Look at some R code files with correlation matrices computed based on vine algorithms.