Introduction to Multivariate Genetic Analysis (2)

Marleen de Moor, Kees-Jan Kan & Nick Martin

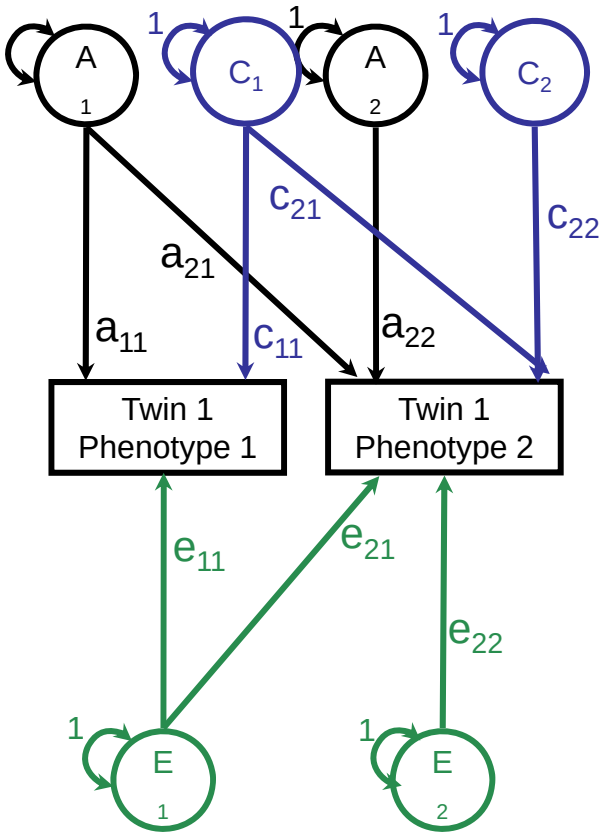
Outline

- 11.00-12.30
  - Lecture Bivariate Cholesky Decomposition
  - Practical Bivariate analysis of IQ and attention problems
- 12.30-13.30 LUNCH
- 13.30-15.00
  - Lecture Multivariate Cholesky Decomposition
  - PCA versus Cholesky
  - Practical Tri- and Four-variate analysis of IQ, educational attainment and attention problems

Outline

- 11.00-12.30
  - Lecture Bivariate Cholesky Decomposition
  - Practical Bivariate analysis of IQ and attention problems
- 12.30-13.30 LUNCH
- 13.30-15.00
  - Lecture Multivariate Cholesky Decomposition
  - PCA versus Cholesky
  - Practical Tri- and Four-variate analysis of IQ, educational attainment and attention problems

Bivariate Cholesky

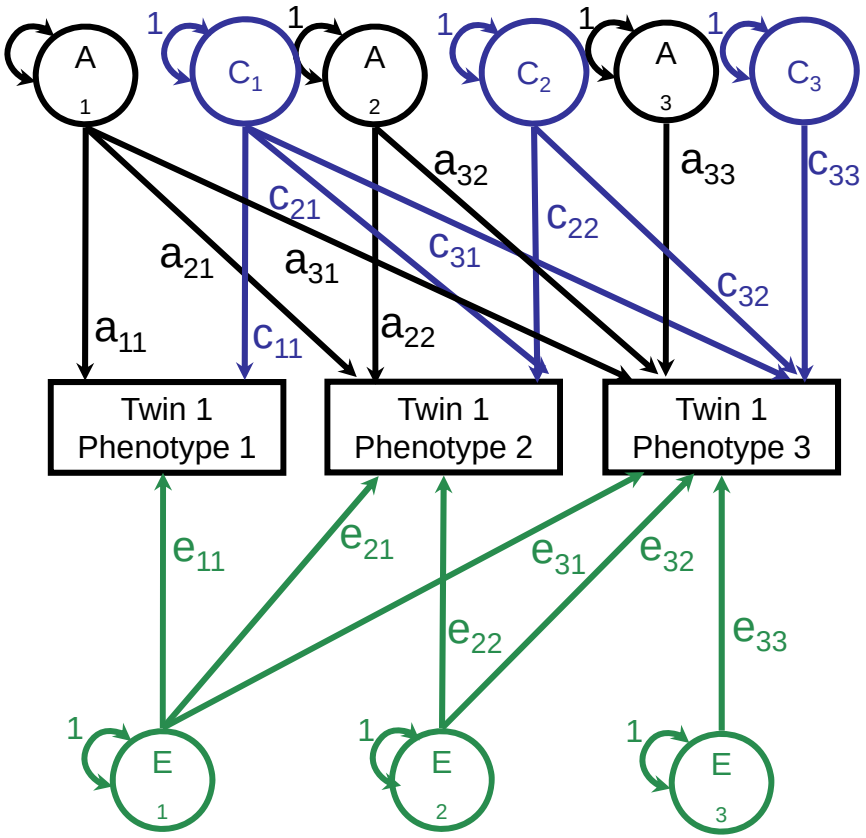


	a1	a2
P1	a_{11}	0
P2		

	c1	c2
P1	c_{11}	0
P2	c_{21}	c_{22}

	e1	e2
P1	e_{11}	0
P2	e_{21}	e_{22}

Adding more phenotypes…

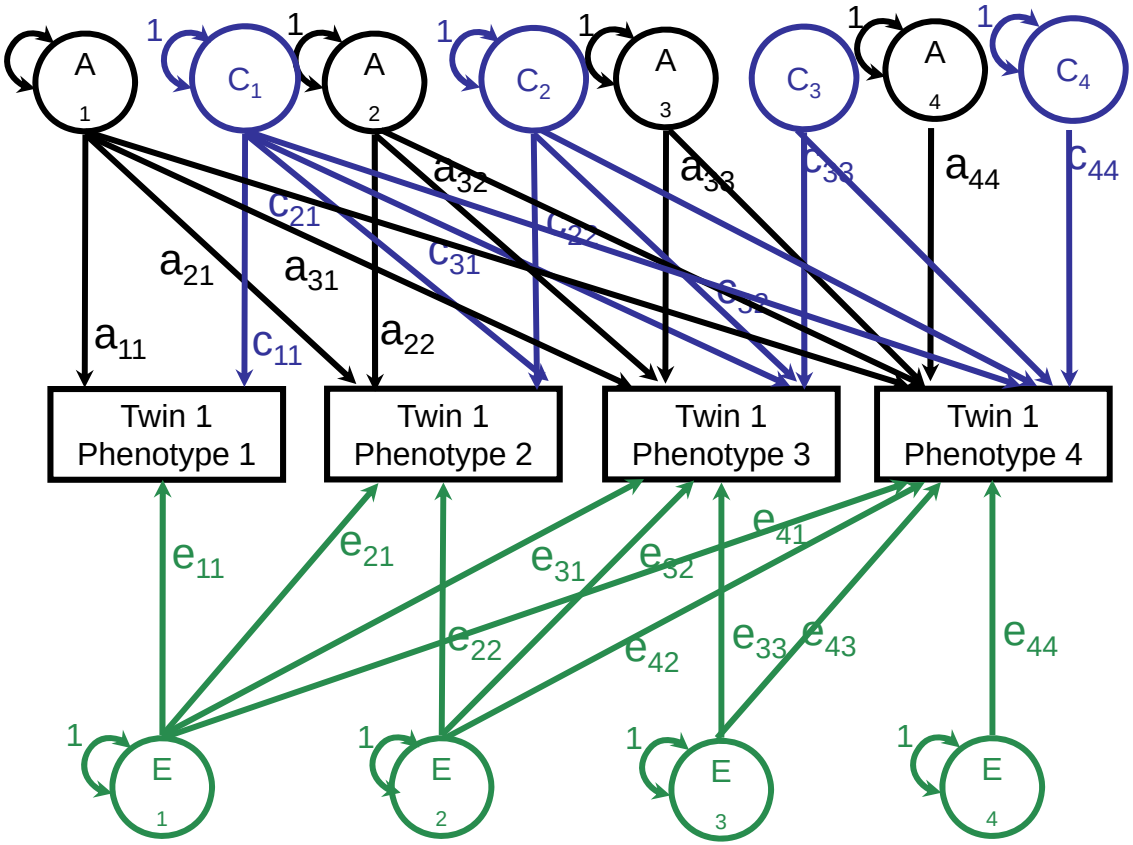


	a1	a2	a3
P1	a_{11}	0	0
P2			0
P3			

	c1	c2	c3
P1	c_{11}	0	0
P2	c_{21}	c_{22}	0
P3	c_{31}	c_{32}	c_{33}

	e1	e2	e3	
P1		e_{11}	0	0
P2		e_{21}	e_{22}	0
P3		e_{31}	e_{32}	e_{33}

Adding more phenotypes…

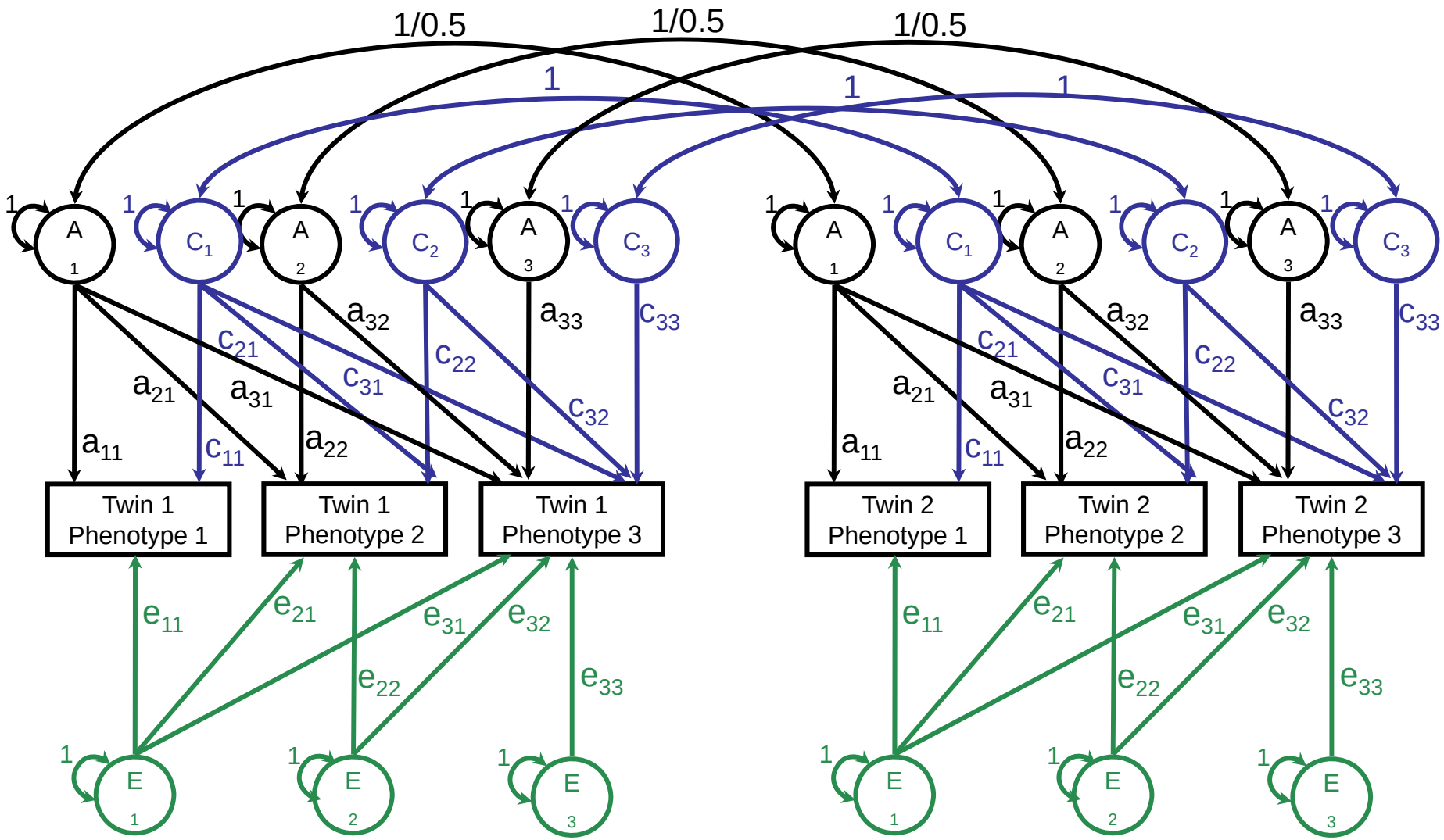


	a1	a2	a3	a4
P1	a_{11}	0	0	0
P2			0	0
P3				0
P4				

	c1	c2	c3	c4
P1	c_{11}	0	0	0
P2	c_{21}	c_{22}	0	0
P3	c_{31}	c_{32}	c_{33}	0
P4	c_{41}	c_{42}	c_{43}	c_{44}

	e1	e2	e3	e4
P1	e_{11}	0	0	0
P2	e_{21}	e_{22}	0	0
P3	e_{31}	e_{32}	e_{33}	0
P4	e_{41}	e_{42}	e_{43}	e_{44}

Trivariate Cholesky



What to change in OpenMx script?

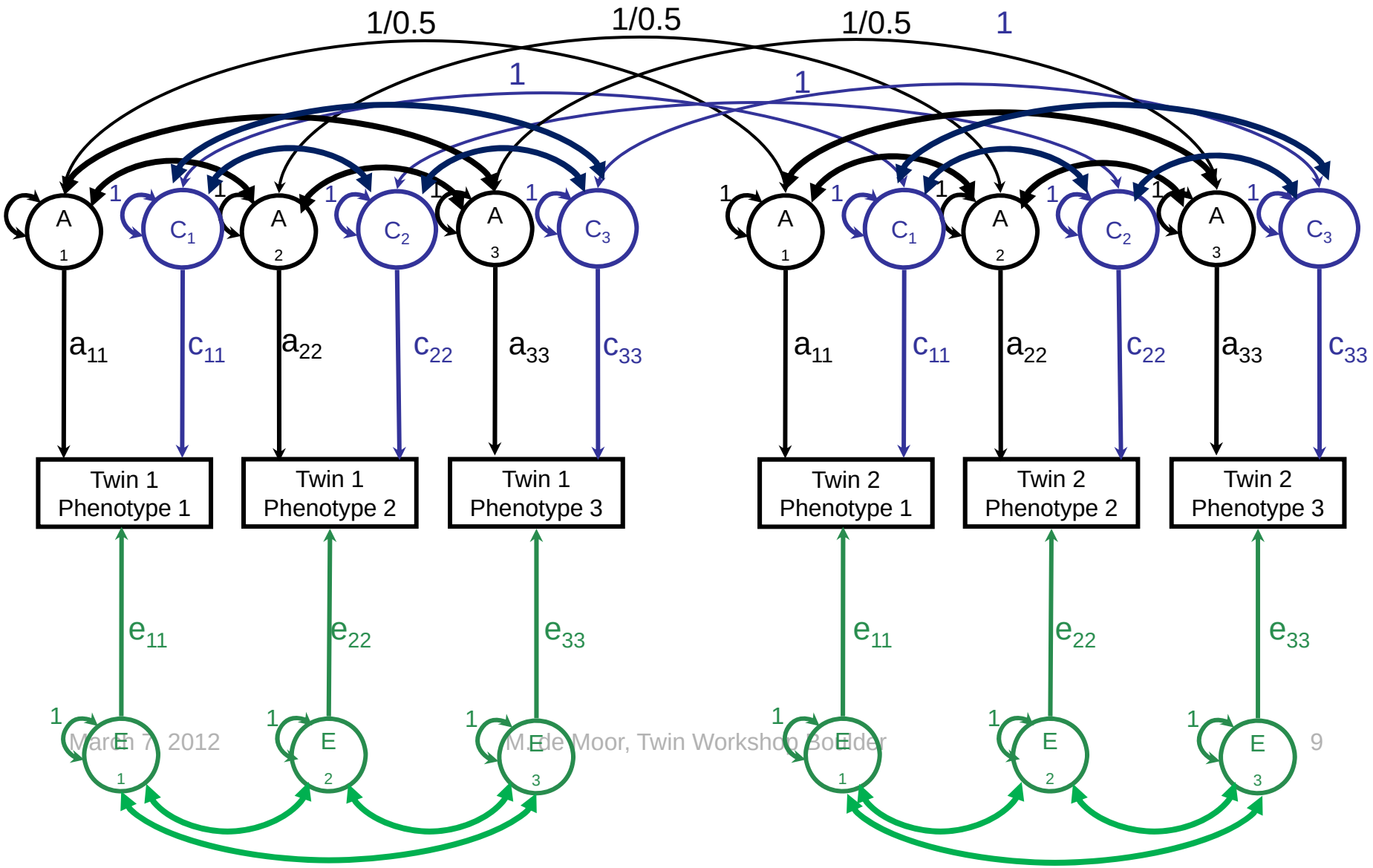
OpenMx

```
Vars <- c('varx', 'vary', 'varz')

nv <- 3
# or, even more efficiently: nv <- length(Vars)
…

# Matrices a, c, and e to store a, c, and e path coefficients
mxMatrix( type="Lower", nrow=nv, ncol=nv, free=TRUE, values=.6, name="a" ),
mxMatrix( type="Lower", nrow=nv, ncol=nv, free=TRUE, values=.6, name="c" ),
mxMatrix( type="Lower", nrow=nv, ncol=nv, free=TRUE, values=.6, name="e" ),
```

Standardized solution – 3 pheno's



Genetic correlations

OpenMx

```
corA <- mxAlgebra(name ="rA", expression = solve(sqrt(I*A))%*%A%*%solve(sqrt(I*A)))
```

2x2

$$\begin{bmatrix} 1 & r_G \\ r_G & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{\sigma^2_{A_{11}}}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma^2_{A_{22}}}} \end{bmatrix} * \begin{bmatrix} \sigma^2_{A_{11}} & \sigma^2_{A_{12}} \\ \sigma^2_{A_{21}} & \sigma^2_{A_{22}} \end{bmatrix} * \begin{bmatrix} \frac{1}{\sqrt{\sigma^2_{A_{11}}}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma^2_{A_{22}}}} \end{bmatrix}$$

3x3

$$\begin{bmatrix} 1 & & \\ r_{G,12} & 1 & \\ r_{G,13} & r_{G,23} & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{a11^2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{a21^2+a22^2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{a31^2+a32^2+a33^2}} \end{bmatrix} * \begin{bmatrix} a11^2 & a11a21 & a11a31 \\ a11a21 & a21^2+a22^2 & a21a31+a22a32 \\ a11a31 & a21a31+a22a32 & a31^2+a32^2+a33^2 \end{bmatrix} * \begin{bmatrix} \frac{1}{\sqrt{a11^2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{a21^2+a22^2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{a31^2+a32^2+a33^2}} \end{bmatrix}$$

$$r_{g,12} = \frac{a_{21}a_{11}}{\sqrt{a_{11}^2 * (a_{21}^2 + a_{22}^2)}} \qquad r_{g,13} = \frac{a_{31}a_{11}}{\sqrt{a_{11}^2 * (a_{31}^2 + a_{32}^2 + a_{33}^2)}} \qquad r_{g,23} = \frac{a_{21}a_{31} + a_{22}a_{32}}{\sqrt{(a_{31}^2 + a_{32}^2 + a_{33}^2) * (a_{21}^2 + a_{22}^2)}}$$

The order of variables

- Order of variables does not matter for the solution!
  - Fit is identical, just different parameterization
  - Standardized solutions are identical in terms of fit and parameter estimates!
- But interpretation of A/C/E variance components is different!
  - Where A2 refers to those genetic factors that are not shared with phenotype 1
- Sometimes there is natural ordering:
  - Temporal ordering (IQ at 2 time points)
  - Neuroticism and MDD symptoms

Cholesky decomposition is not a model…

- No constraints on covariance matrices
- Just reparameterization…
- …But very useful to explore the data!
- Observed statistics = Number of parameters

Cholesky decomposition is not a model…

- Bivariate constrained saturated model:
  - 2 variances, 1 within-twin covariance MZ=DZ
  - 2 within-trait cross-twin covariances MZ
  - 1 cross-trait cross-twin covariance MZ
  - 2 within-trait cross-twin covariances DZ
  - 1 cross-trait cross-twin covariance DZ

9 observed statistics

- Bivariate Cholesky decomposition
  - a11, a21, a22
  - c11, c21, c22
  - e11, e21, e22

9 parameters

Comparison with other models

Cholesky decomposition models

Principal component analysis
→ Sanja, now

Genetic factor models
→ Hermine, after coffee break

Confirmatory factor models
→ Dorret, Sanja, Michel, this morning

Further reading

Three classic papers:

- Martin NG, Eaves LJ: The genetical analysis of covariance structure. *Heredity* 38:79-95, 1977
- Carey, G. Inference About Genetic Correlations, BG, 1988
- Loehlin, J. The Cholesky Approach: A Cautionary Note, BG, 1996
- Carey, G. Cholesky Problems, BG, 2005

SEE ALSO:

http://genepi.qimr.edu.au/staff/classicpapers/

Outline

- 11.00-12.30
  - Lecture Bivariate Cholesky Decomposition
  - Practical Bivariate analysis of IQ and attention problems
- 12.30-13.30 LUNCH
- 13.30-15.00
  - Lecture Multivariate Cholesky Decomposition
  - PCA versus Cholesky
  - Practical Tri- and Four-variate analysis of IQ, educational attainment and attention problems

Practical

- Trivariate ACE Cholesky model

- 126 MZ and 126 DZ twin pairs from Netherlands Twin Register
- Age 12

- Educational achievement (EA)
- FSIQ
- Attention Problems (AP) [mother-report]

Practical

- Script CholeskyTrivariate.R
- Dataset Cholesky.dat

Exercise

- Add Educational Achievement as the first of the 3 variables
- Run the saturated model, ACE model and AE model
- Question: Can we drop C?

	-2LL	df	chi2	Δdf	P-value
ACE model			-	-	-
AE model					

Exercise

- Run 4 submodels
  - Submodel 1: drop rg between EA and AP
  - Submodel 2: drop rg between FSIQ and AP
  - Submodel 3: drop re between EA and AP
  - Submodel 4: drop re between FSIQ and AP
- Compare fit of each submodel with full AE model

Exercise

- Questions:
  - Can we drop rg between EA and AP?
  - Can we drop rg between FSIQ and AP?
  - Can we drop re between EA and AP?
  - Can we drop re between FSIQ and AP?

	-2LL	df	chi2	Δdf	P-value
AE model			-	-	-
No a31					
No a32					
No e31					
No e32					

Extra exercise

- Replace FSIQ by VIQ and PIQ, and run a fourvariate Cholesky model.

- Questions:
  - Is AP differentially related to VIQ and PIQ, phenotypically and genotypically?