Simplex Model
Practical

Conor Dolan, Dorret Boomsma & Eveline de Zeeuw

Thanks to many others

COPY FILES FROM:

/home/eveline/2020/Simplex

Example

- Measurement
  - Based on mother-reported height and weight
  - BMI at ages 4, 7, 10 and 12
  - Only complete cases
  - Residualized for sex and birth cohort of the child
- Sample
  - Netherlands Twin Register
  - 843 MZ and 1312 DZ pairs
- Research question
  - Do genes and/or the environment explain stability and change in BMI during childhood?

Models

1) Saturated Model

2) ACE Model

MZ	=	A+C+E	A+C		
		A+C		A+C+E	
DZ	=	A+C+E	.5*A+C		
		.5*A+C		A+C+E	

3) Simplex Model

MZ	=	A+C+E	A+C		$A = (I-\beta_A)\ \Psi_A\ (I-\beta_A)^t + \Theta_A$
		A+C		A+C+E	$C = (I-\beta_C)\ \Psi_C\ (I-\beta_C)^t + \Theta_C$
DZ	=	A+C+E	.5*A+C		$E = (I-\beta_E)\ \Psi_E\ (I-\beta_E)^t + \Theta_E$
		.5*A+C	A+C+E		

Saturated Model

	Age 4 **T1**	Age 7 **T1**	Age 10 **T1**	Age 12 **T1**	Age 4 **T2**	Age 7 **T2**	Age 10 **T2**	Age 12 **T2**
Age 4 **T1**	var1							
Age 7 **T1**	cov21	var2						
Age 10 **T1**	cov31	cov32	var3					
Age 12 **T1**	cov41	cov42	cov43	var4				
Age 4 **T2**	cov51	cov52	cov53	cov54	var5			
Age 7 **T2**	cov61	cov62	cov63	cov64	cov65	var6		
Age 10 **T2**	cov71	cov72	cov73	cov74	cov75	cov76	var7	
Age 12 **T2**	cov81	cov82	cov83	cov84	cov85	cov86	cov87	var8

MZ ≠ DZ

Age 4 **T1**	Age 7 **T1**	Age 10 **T1**	Age 12 **T1**	Age 4 **T2**	Age 7 **T2**	Age 10 **T2**	Age 12 **T2**
mean1	mean2	mean3	mean4	mean1	mean2	mean3	mean4

MZ = DZ

Correlations

Within-twin cross-time point

	Age 4 **T1**	Age 7 **T1**	Age 10 **T1**	Age 12 **T1**	Age 4 **T2**	Age 7 **T2**	Age 10 **T2**	Age 12 **T2**
Age 4 **T1**	1	0.593	0.499	0.432	0.448	0.250	0.208	0.179
Age 7 **T1**	0.591	1	0.738	0.669	0.295	0.511	0.370	0.346
Age 10 **T1**	0.543	0.762	1	0.760	0.239	0.367	0.469	0.400
Age 12 **T1**	0.487	0.683	0.819	1	0.23	0.351	0.403	0.484
Age 4 **T2**	0.802	0.504	0.466	0.405	1	0.592	0.508	0.450
Age 7 **T2**	0.529	0.864	0.686	0.626	0.601	1	0.732	0.676
Age 10 **T2**	0.491	0.689	0.879	0.757	0.538	0.760	1	0.811
Age 12 **T2**	0.438	0.624	0.757	0.902	0.453	0.688	0.829	1

4-7	4-10	4-12	7-10	7-12	10-12
.59-.60	.50-.54	.43-.49	.73-.76	.67-.69	.76-.83

Correlations

Cross-twin within-time point

	Age 4 **T1**	Age 7 **T1**	Age 10 **T1**	Age 12 **T1**	Age 4 **T2**	Age 7 **T2**	Age 10 **T2**	Age 12 **T2**
Age 4 **T1**	1	0.593	0.499	0.432	0.448	0.250	0.208	0.179
Age 7 **T1**	0.591	1	0.738	0.669	0.295	0.511	0.370	0.346
Age 10 **T1**	0.543	0.762	1	0.760	0.239	0.367	0.469	0.400
Age 12 **T1**	0.487	0.683	0.819	1	0.230	0.351	0.403	0.484
Age 4 **T2**	0.802	0.504	0.466	0.405	1	0.592	0.508	0.450
Age 7 **T2**	0.529	0.864	0.686	0.626	0.601	1	0.732	0.676
Age 10 **T2**	0.491	0.689	0.879	0.757	0.538	0.760	1	0.811
Age 12 **T2**	0.438	0.624	0.757	0.902	0.453	0.688	0.829	1

	Age 4	Age 7	Age 10	Age 12
MZ	.80	.86	.88	.90
DZ	.45	.51	.47	.48

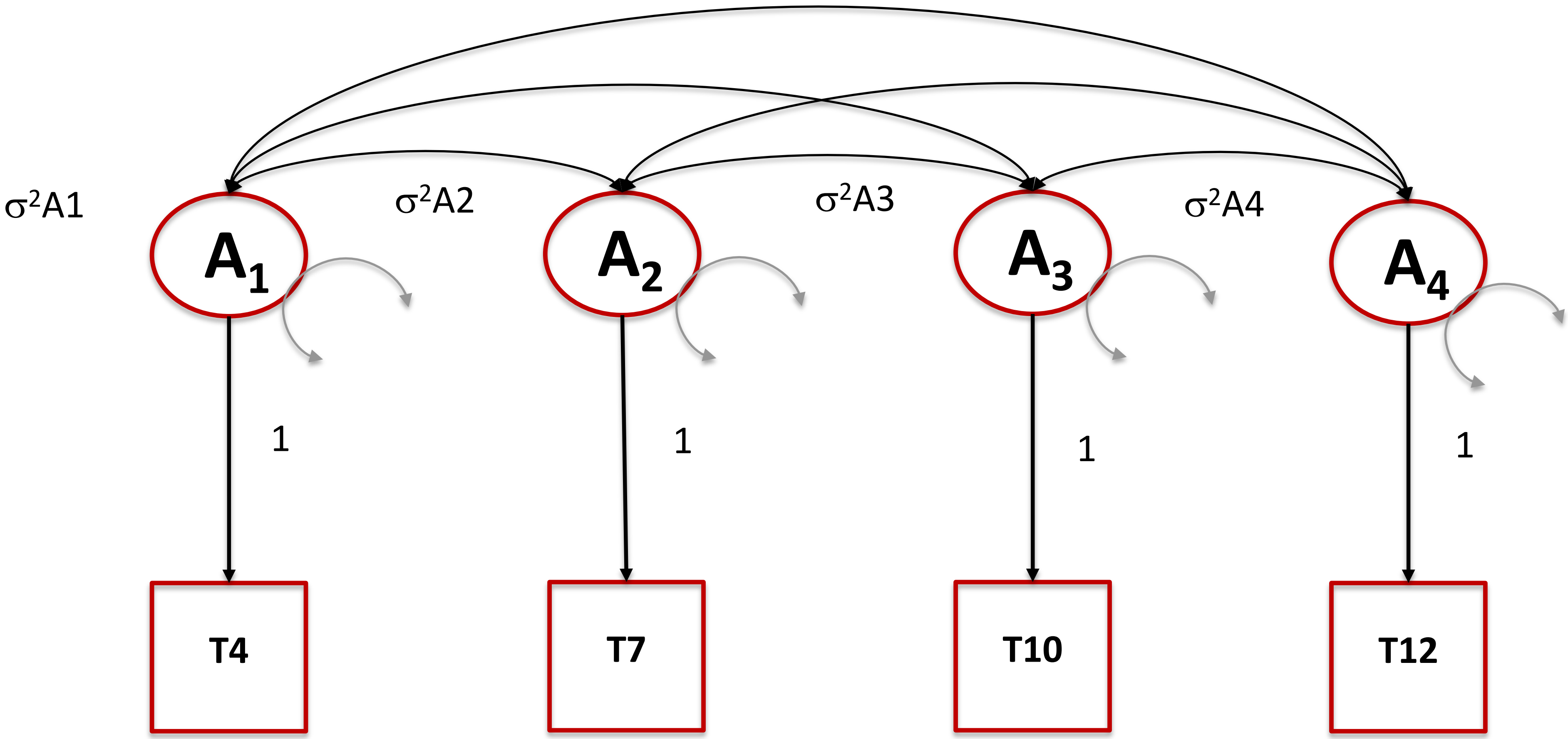
Correlations

Cross-twin cross-time point

	Age 4 **T1**	Age 7 **T1**	Age 10 **T1**	Age 12 **T1**	Age 4 **T2**	Age 7 **T2**	Age 10 **T2**	Age 12 **T2**
Age 4 **T1**	1	0.593	0.499	0.432	0.448	0.250	0.208	0.179
Age 7 **T1**	0.591	1	0.738	0.669	0.295	0.511	0.370	0.346
Age 10 **T1**	0.543	0.762	1	0.760	0.239	0.367	0.469	0.400
Age 12 **T1**	0.487	0.683	0.819	1	0.230	0.351	0.403	0.484
Age 4 **T2**	0.802	0.504	0.466	0.405	1	0.592	0.508	0.450
Age 7 **T2**	0.529	0.864	0.686	0.626	0.601	1	0.732	0.676
Age 10 **T2**	0.491	0.689	0.879	0.757	0.538	0.760	1	0.811
Age 12 **T2**	0.438	0.624	0.757	0.902	0.453	0.688	0.829	1

	4-7	4-10	4-12	7-10	7-12	10-12
MZ	.50-.53	.47-.49	.41-.44	.69-.69	.62-.63	.76-.76
DZ	.25-.29	.21-.24	.18-.23	.37-.37	.35-.35	.40-.40

Multivariate ACE



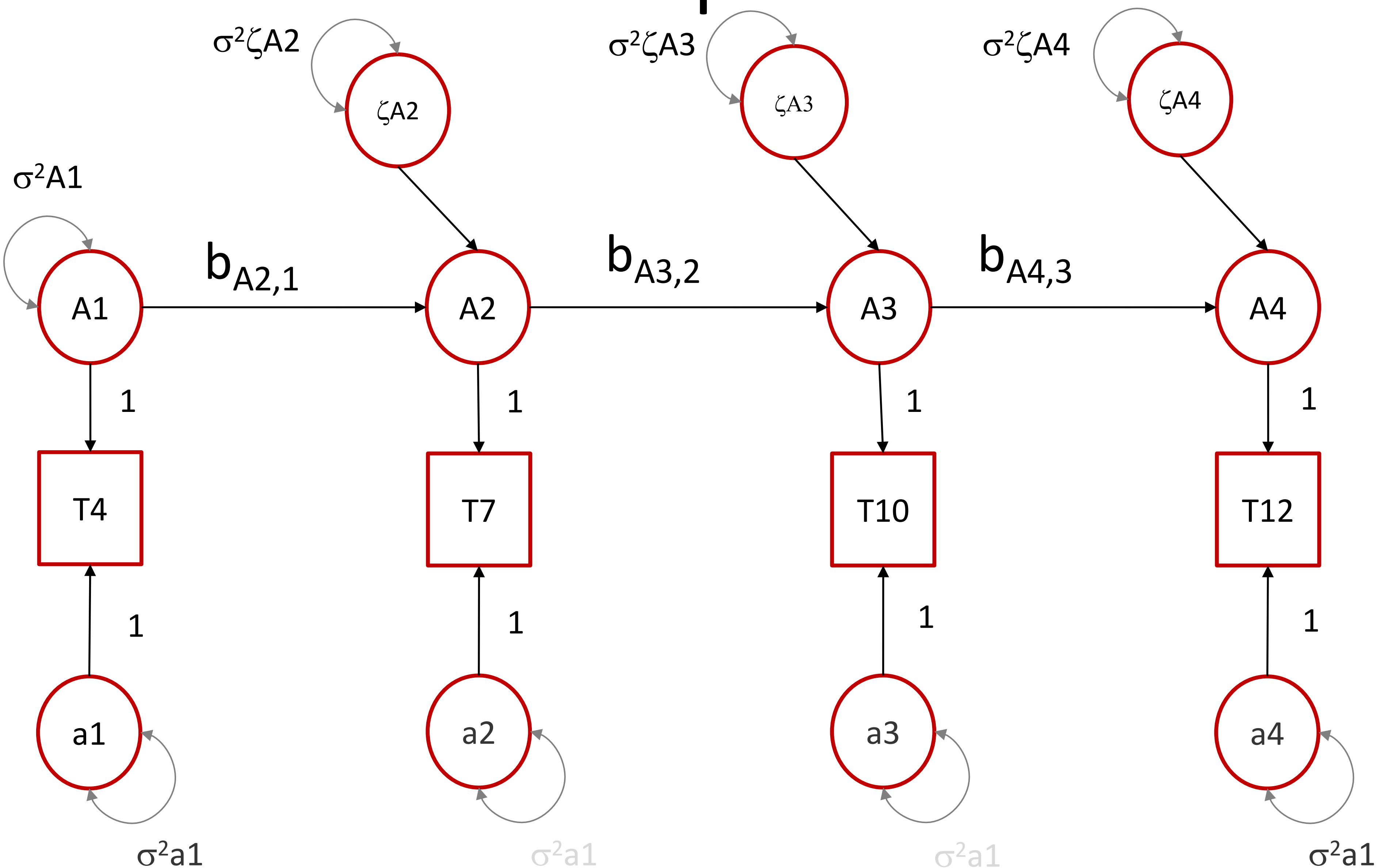
ACE

		Age 4	Age 7	Age 10	Age 12
A	Age 4	0.680	0.780	0.841	0.829
	Age 7	0.780	0.669	0.767	0.760
	Age 10	0.841	0.767	0.801	0.793
	Age 12	0.829	0.760	0.793	0.822
C	Age 4	0.114	0.078	0.036	0.060
	Age 7	0.078	0.186	0.129	0.147
	Age 10	0.036	0.129	0.074	0.121
	Age 12	0.060	0.147	0.121	0.078
E	Age 4	0.205	0.142	0.123	0.110
	Age 7	0.142	0.145	0.104	0.093
	Age 10	0.123	0.104	0.125	0.086
	Age 12	0.110	0.093	0.086	0.100

ACE

		Age 4	Age 7	Age 10	Age 12
A	Age 4	1	0.686	0.587	0.499
	Age 7	0.686	1	0.779	0.693
	Age 10	0.587	0.779	1	0.780
	Age 12	0.499	0.693	0.780	1
C	Age 4	1	0.317	0.202	0.289
	Age 7	0.317	1	0.816	0.826
	Age 10	0.202	0.816	1	1.276
	Age 12	0.289	0.826	1.276	1
E	Age 4	1	0.487	0.396	0.345
	Age 7	0.487	1	0.574	0.522
	Age 10	0.396	0.574	1	0.611
	Age 12	0.345	0.522	0.611	1

Simplex



σ²ζA2
ζA2
σ²ζA3
ζA3
σ²ζA4
ζA4
σ²A1
A1
$b_{A2,1}$
A2
$b_{A3,2}$
A3
$b_{A4,3}$
A4
1
1
1
1
T4
T7
T10
T12
1
1
1
1
a1
a2
a3
a4
σ²a1
σ²a1
σ²a1
σ²a1

Data

```
# Load Data
BMIdata    <-  read.table(file='longdataBMI.dat',header=TRUE, na=-999)
head(BMIdata)
dim(BMIdata)

# Select Variables for Analysis
vars      <- c('BMI_age4','BMI_age7','BMI_age10','BMI_age12') # list of variables names
nv        <- 4                                                 # number of variables
ntv       <- nv*2                                              # number of total variables
nt        <- 4                                                 # number of time points
selVars   <- paste(vars,c(rep('_tw1',nv),rep('_tw2',nv)),sep="")

# Select Data for Analysis
mzData    <- subset(BMIdata, zyg==1|zyg==3, selVars)
dzData    <- subset(BMIdata, zyg==2|zyg==4|zyg==5|zyg==6, selVars)

# Generate Descriptive Statistics
round(colMeans(mzData,na.rm=TRUE),3)
round(colMeans(dzData,na.rm=TRUE),3)
round(cov(mzData,use="complete"),3)
round(cov(dzData,use="complete"),3)
round(cor(mzData,use="complete"),3)
round(cor(dzData,use="complete"),3)
```

Means

```
# Create Algebra for expected Mean Matrices
expMean <- mxMatrix(type='Full',nrow=1,ncol=ntv,free=T,
                    labels=c('m4','m7','m10','m12','m4','m7','m10','m12'),
                    value=.1,name='expMean')
```

Variance Components

```
# Create Algebra for Variance Components
PsA   <- mxMatrix(type='Diag',nrow=nv,ncol=nv,free=c(T,T,T,T),labels=c("PsA11","PsA22","PsA33","PsA44"),value=c(1,1,1,1),name='PsA')
PsC   <- mxMatrix(type='Diag',nrow=nv,ncol=nv,free=c(T,T,T,T),labels=c("PsC11","PsC22","PsC33","PsC44"),value=c(1,1,1,1),name='PsC')
PsE   <- mxMatrix(type='Diag',nrow=nv,ncol=nv,free=c(T,T,T,T),labels=c("PsE11","PsE22","PsE33","PsE44"),value=c(1,1,1,1),name='PsE')
```

```
# Create Algebra for Variance Components
TeA   <- mxMatrix(type='Diag',nrow=nv,ncol=nv,free=c(T,T,T,T),labels=c("TeA11","TeA11","TeA11","TeA11"),value=c(.1,.1,.1,.1),name='TeA')
TeC   <- mxMatrix(type='Diag',nrow=nv,ncol=nv,free=c(T,T,T,T),labels=c("TeC11","TeC11","TeC11","TeC11"),value=c(.1,.1,.1,.1),name='TeC')
TeE   <- mxMatrix(type='Diag',nrow=nv,ncol=nv,free=c(T,T,T,T),labels=c("TeE11","TeE11","TeE11","TeE11"),value=c(.1,.1,.1,.1),name='TeE')
```

Variance Components

```
# Create Algebra for Beta's
BeA    <- mxMatrix(type='Full',nrow=nv,ncol=nv,
         free=matrix(c(F,F,F,F,
                       T,F,F,F,
                       F,T,F,F,
                       F,F,T,F),nv,nv,byrow=T),
         labels=matrix(c('NA','NA','NA','NA',
                         'BeA21','NA','NA','NA',
                         'NA','BeA32','NA','NA',
                         'NA','NA','BeA43','NA'),nv,nv,byrow=T),
         value=matrix(c(0,0,0,0,
                        .5,0,0,0,
                        0,.5,0,0,
                        0,0,.5,0),nv,nv,byrow=T),name='BeA')
BeC    <- mxMatrix(type='Full',nrow=nv,ncol=nv,
                  free=matrix(c(F,F,F,F,
                                T,F,F,F,
                                F,T,F,F,
                                F,F,T,F),nv,nv,byrow=T),
                  labels=matrix(c('NA','NA','NA','NA',
                                  'BeC21','NA','NA','NA',
                                  'NA','BeC32','NA','NA',
                                  'NA','NA','BeC43','NA'),nv,nv,byrow=T),
                  value=matrix(c(0,0,0,0,
                                 .5,0,0,0,
                                 0,.5,0,0,
                                 0,0,.5,0),nv,nv,byrow=T),name='BeC')
BeE    <- mxMatrix(type='Full',nrow=nv,ncol=nv,
                  free=matrix(c(F,F,F,F,
                                T,F,F,F,
                                F,T,F,F,
                                F,F,T,F),nv,nv,byrow=T),
                  labels=matrix(c('NA','NA','NA','NA',
                                  'BeE21','NA','NA','NA',
                                  'NA','BeE32','NA','NA',
                                  'NA','NA','BeE43','NA'),nv,nv,byrow=T),
                  value=matrix(c(0,0,0,0,
                                 .5,0,0,0,
                                 0,.5,0,0,
                                 0,0,.5,0),nv,nv,byrow=T),name='BeE')
```

$$A = (I-\beta_A)\ \Psi_A\ (I-\beta_A)^t + \Theta_A$$

$$C = (I-\beta_C)\ \Psi_C\ (I-\beta_C)^t + \Theta_C$$

$$E = (I-\beta_E)\ \Psi_E\ (I-\beta_E)^t + \Theta_E$$

```
# Create Algebra for Covariances
Imat    <- mxMatrix(type='Iden',nrow=nv,ncol=nv,name='I')
IBeA    <- mxAlgebra(expression=solve(I-BeA),name='iBeA')
IBeC    <- mxAlgebra(expression=solve(I-BeC),name='iBeC')
IBeE    <- mxAlgebra(expression=solve(I-BeE),name='iBeE')
covA    <- mxAlgebra(expression=iBeA%*%(PsA)%*%t(iBeA)+(TeA),name='A')
covC    <- mxAlgebra(expression=iBeC%*%(PsC)%*%t(iBeC)+(TeC),name='C')
covE    <- mxAlgebra(expression=iBeE%*%(PsE)%*%t(iBeE)+(TeE),name='E')
```

Covariances

```
# Create Algebra for expected Variance/Covariance Matrices in MZ and DZ twins
expCovMZ  <- mxAlgebra( expression= rbind(cbind(A+C+E, A+C),
                                          cbind(A+C, A+C+E)), name="expCovMZ")
expCovDZ  <- mxAlgebra( expression= rbind(cbind(A+C+E, 0.5%x%A+C),
                                          cbind(0.5%x%A+C, A+C+E)), name="expCovDZ")
```

```r
# Create Data Objects for Multiple Groups
dataMZ   <- mxData(observed=mzData, type="raw")
dataDZ   <- mxData(observed=dzData, type="raw")

# Create Expectation Objects for Multiple Groups
expMZ    <- mxExpectationNormal(covariance="expCovMZ", means="expMean", dimnames=selVars)
expDZ    <- mxExpectationNormal(covariance="expCovDZ", means="expMean", dimnames=selVars)
funML    <- mxFitFunctionML()

# Create Model Objects for Multiple Groups
pars     <- list(covA, covC, covE, Imat, IBeA, IBeC, IBeE, BeA, BeC, BeE, TeA, TeC, TeE, PsA, PsC, PsE,
                 covP, covSA, covSC, covSE, corA, corC, corE)
modelMZ  <- mxModel(name="MZ", pars, expMean, expCovMZ, dataMZ, expMZ, funML)
modelDZ  <- mxModel(name="DZ", pars, expMean, expCovDZ, dataDZ, expDZ, funML)
multi    <- mxFitFunctionMultigroup(c("MZ","DZ"))

# Build Model
simplexACEModel  <- mxModel("simplexACE", modelMZ, modelDZ, multi )

# Run Model
simplexACEFit    <- mxTryHard(simplexACEModel)
(simplexACESumm  <- summary(simplexACEFit))
mxCompare(satFit, simplexACEFit)
```

Practical

1. Open the script *simplexModel.R*
2. Walk through the first part of the script

   - 1 = MZm, 2 = DZm, 3 = MZf, 4 = DZ4, 5 = DZmf, 6 = DZfm

1. Run the first part of the script
2. Run the submodels ($\alpha = 0.01$)

   - For each model fill out the question marks

1. Make a table with the model fits
2. Report the output of the best model
3. Make sure that you know what you are doing

Simplex

	A				C				E			
Ψ	0.577	0	0	0	0.131	0	0	0	0.162	0	0	0
	0	0.230	0	0	0	0.143	0	0	0	0.052	0	0
	0	0	0.174	0	0	0	0.035	0	0	0	0.025	0
	0	0	0	0.193	0	0	0	-0.032	0	0	0	0.005
β	0	0	0	0	0	0	0	0	0	0	0	0
	0.718	0	0	0	0.434	0	0	0	0.482	0	0	0
	0	0.985	0	0	0	0.560	0	0	0	0.792	0	0
	0	0	0.870	0	0	0	1.206	0	0	0	0.816	0
Θ	0.066	0	0	0	-0.003	0	0	0	0.038	0	0	0
	0	0.066	0	0	0	-0.003	0	0	0	0.038	0	0
	0	0	0.066	0	0	0	-0.003	0	0	0	0.038	0
	0	0	0	0.066	0	0	0	-0.003	0	0	0	0.038

Results

Model		EP	-2LL	df	AIC	Δ-2LL	Δdf	P
1. SAT		76	32071.71	16712	1352.29	-	-	-
2. ACE	1	34	32135.91	16754	1372.09	64.20	42	.015
3. Simplex ACE	1	28	32141.94	16760	1378.06	70.23	48	.020
4. No Neg Vars	3	26	32146.57	16762	1377.43	4.63	1	.099
5. Simplex AE	3	20	32196.65	16768	1339.35	50.09	6	<.001
7. No Θ_A	6	25	32171.44	16763	1354.56	24.87	1	<.001
8. Only Ψ_E at T1	6	23	32369.99	16765	1160.01	223.42	1	<.001

Results

	A				C				E			
Ψ	0.579	0	0	0	0.127	0	0	0	0.162	0	0	0
	0	0.230	0	0	0	0.142	0	0	0	0.052	0	0
	0	0	0.179	0	0	0	0.030	0	0	0	0.024	0
	0	0	0	0.165	0	0	0	0	0	0	0	0.006
β	0	0	0	0	0	0	0	0	0	0	0	0
	0.715	0	0	0	0.450	0	0	0	0.483	0	0	0
	0	0.979	0	0	0	0.59	0	0	0	0.798	0	0
	0	0	0.886	0	0	0	1.096	0	0	0	0.808	0
Θ	0.064	0	0	0	0	0	0	0	0.039	0	0	0
	0	0.064	0	0	0	0	0	0	0	0.039	0	0
	0	0	0.064	0	0	0	0	0	0	0	0.039	0
	0	0	0	0.064	0	0	0	0	0	0	0	0.039

AIC weights

- Inference of the model
  - Choosing the ‘right’ order to drop parameters is complex
  - Choosing whether to drop parameters at all
- One solution is to not choose a best model, but to use AIC weights (Wagemakers & Farrell, 2004)
- AIC weights represent the relative probability a model is closest to the ‘true’ model given all the models you consider

Questions?