

(Re)introduction to Mx

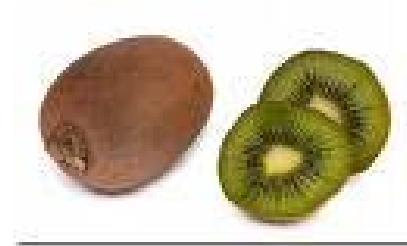
Sarah Medland



Kiwi



Chinese Gooseberry





Starting at the beginning

- Data preparation
 - Mx expects 1 line per case/family
 - Almost limitless number of families and variables
 - Space delimited is best
 - Can use a missing code ie -9 or can use the default '.'



Important structural stuff

- Script is composed of one or more jobs (can handle many 'nested' jobs in one script or 2 non nested jobs)
- Each job is composed of one or more groups
 - Each group is 'opened' with a title
 - Each group is 'closed' with an end statement
 - You must tell Mx how many groups will be in the job



A bit about groups

- 3 types of groups
 - Calculation
 - Data
 - If analysing raw data Mx expects a Means Model and a Covariance Model
 - Constraint



Matrices: the building blocks

- Many types
- Denoted by a single letter
 - Elements defined by letter and 3 numbers
 - A_{121} = A matrix group 1 row 2 column 1
- All constants and estimated parameters must be placed in a matrix & M_x must be told what type of matrix it is
- Letters can be reused in subsequent groups

Matrices: the building blocks

- Many types

A Zero 2 3 Free

```
0 0 0
0 0 0
```

B Unit 2 3 Free

```
0 0 0
0 0 0
```

C Iden 3 3 Free

```
1 0 0
0 1 0
0 0 1
```

F Diag 3 3 Free

```
1 0 0
0 2 0
0 0 3
```

```
? 0 0
0 ? 0
0 0 ?
```

G Sdiag 3 3 Free

```
0 0 0
1 0 0
2 3 0
```

```
0 0 0
? 0 0
? ? 0
```

H Stand 3 3 Free

```
0 1 2
1 0 3
2 3 0
```

```
1 ? ?
? 1 ?
? ? 1
```

I Symm 3 3 Free

```
1 2 4
2 3 5
4 5 6
```

```
? ? ?
? ? ?
? ? ?
```

J Lower 3 3 Free

```
1 0 0
2 3 0
4 5 6
```

```
? 0 0
? ? 0
? ? ?
```

K Full 2 4 Free

```
1 2 3 4
5 6 7 8
```

```
? ? ? ?
? ? ? ?
```



Short cuts

- Anything after ! is read as a comment
- Can predefine frequently used/changed parameters
 - `#define nvar2=2`
- Can read in another file within the script
 - `#include ozbmi2.dat`
- Can run loops – via the repeat command
- Use an end of line signal (; or /) except in the Labels command



Setting up the script – calculation group

- 1st line is the title
- 2nd specifies group type
- Matrix definition
 - Begin Matrices – End Matrices
 - If a matrix is not specified free it will be considered fixed
- Algebra
 - Begin Algebra – End Algebra
- Starting values for free/estimated parameters or specified values for constants
- End



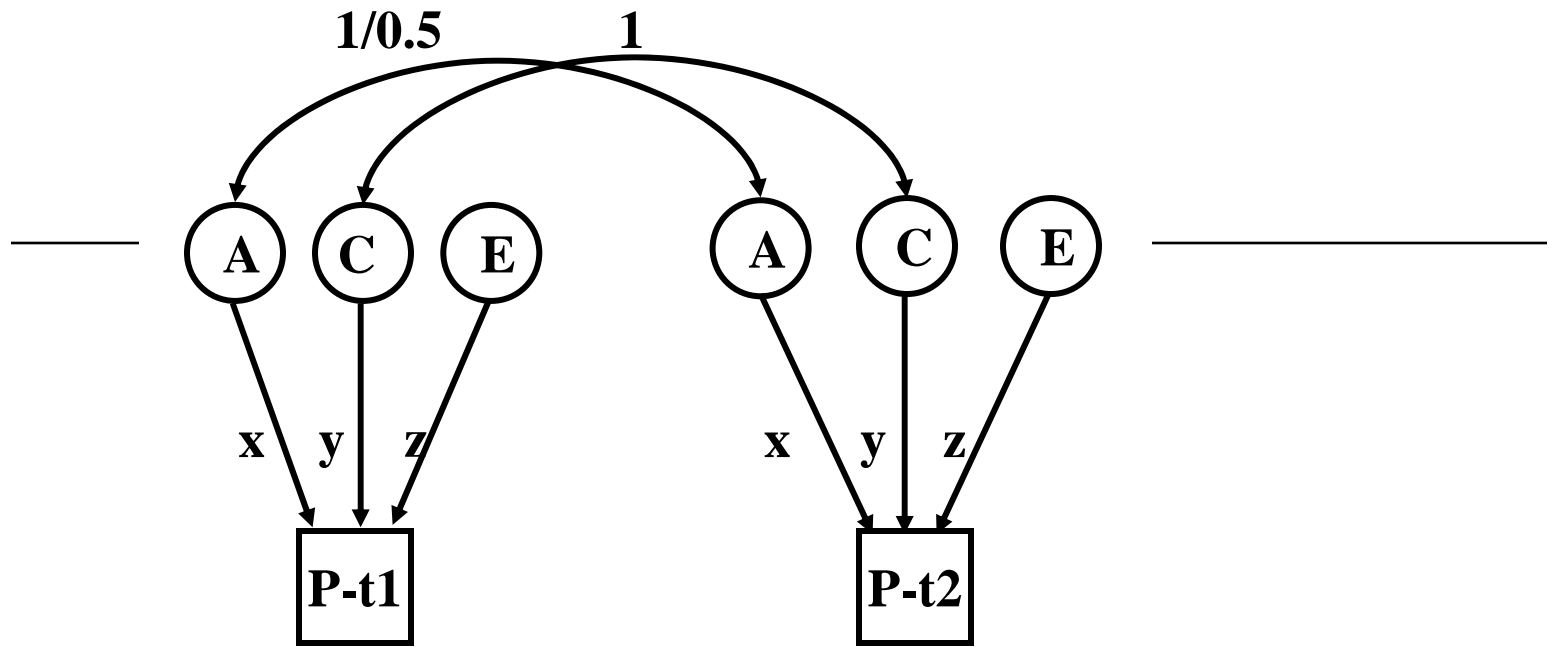
Setting up the script – data group

- 1st line is the title
- 2nd specifies group type and number of variables (in ozbmi2.dat)
- 3rd line gives data location (in ozbmi2.dat)
 - Rectangular file = continuous data
 - Ordinal file = ordinal data (Mx will expect a thresholds model not a means model)
- List the variables (in ozbmi2.dat)
- Select if ...
- Select variables
 - Order is important! Select all vs for twin1 then twin2 then sib1 ect
- Specify which vs are covariates (definition variables)



Setting up the script – data group

- Matrix definition
 - Call matrices from previous groups and/or define new matrices
- Algebra & starting values
- Means Model
 - can include covariates ie age, sex ...
- Covariance Model
 - Expected to be $nsib * nvar$ by $nsib * nvar$
- End



Variance/covariance matrices

MZ

DZ

	t1	t2
t1	$a^2+c^2+e^2$	a^2+c^2
t2	a^2+c^2	$a^2+c^2+e^2$

	t1	t2
t1	$a^2+c^2+e^2$	$0.5a^2+c^2$
t2	$0.5a^2+c^2$	$a^2+c^2+e^2$



So what do you get

- Mx starts by reading back the script

```
** Mx startup successful **
```

```
**MX-PC 1.63h** Job started on 03/07/06 at 01:51:09
```

```
? ESTIMATE GENETIC AND ENVIRONMENTAL COMPONENTS - ACED MODEL  
? OZ BMI DATA - YOUNGER FEMALES
```

```
The following MX script lines were read for group    1
```

```
#NGROUPS 4
```

```
Note: #NGroup set number of groups to 4
```

```
#DEFINE NVAR 1
```

```
#DEFINE NVAR2 2
```

```
TITLE G1: MODEL PARAMETERS
```

```
CALCULATION
```

```
BEGIN MATRICES;
```

```
X LOWER NVAR NVAR FREE      ? ADDITIVE GENETIC PATH, A  
Y LOWER NVAR NVAR           ? COMMON ENVIRONMENTAL PATH, C  
Z LOWER NVAR NVAR FREE      ? SPECIFIC ENVIRONMENTAL PATH, E  
W LOWER NVAR NVAR FREE      ? DOMINANCE GENETIC PATH, D
```



So what do you get

- Data summary

Summary of UL file data for group 2

	BMI1	BMI2
Code	1.0000	2.0000
Number	647.0000	651.0000
Mean	21.9076	21.8684
Variance	0.9801	0.9528
Minimum	19.4150	18.5724
Maximum	26.0885	25.9592

Summary of UL file data for group 3

	BMI1	BMI2
Code	1.0000	2.0000
Number	389.0000	391.0000
Mean	21.8207	21.8615
Variance	0.9155	1.0274
Minimum	19.4145	19.6650
Maximum	26.1492	25.6367

So what do you get

- Parameter specifications

PARAMETER SPECIFICATIONS

GROUP NUMBER: 1

Title G1: Model Parameters

MATRIX A

This is a computed FULL matrix of order 1 by 1
It has no free parameters specified

MATRIX C

This is a computed FULL matrix of order 1 by 1
It has no free parameters specified

MATRIX D

This is a computed FULL matrix of order 1 by 1
It has no free parameters specified

```
MATRIX W
This is a LOWER TRIANGULAR matrix of order 1 by 1
1
DOM_GEN 3
```

```
MATRIX X
This is a LOWER TRIANGULAR matrix of order 1 by 1
1
ADD_GEN 1
```

```
MATRIX Y
This is a LOWER TRIANGULAR matrix of order 1 by 1
It has no free parameters specified
```

```
MATRIX Z
This is a LOWER TRIANGULAR matrix of order 1 by 1
1
SPEC_ENU 2
```

So what do you get

- Estimates

```
Mx starting optimization; number of parameters = 7
```

```
MX PARAMETER ESTIMATES
```

```
GROUP NUMBER: 1
```

```
Title G1: Model Parameters
```

```
MATRIX A
```

```
This is a computed FULL matrix of order 1 by 1  
[=X*X']
```

```
1  
1 0.5613
```

```
MATRIX C
```

```
This is a computed FULL matrix of order 1 by 1  
[=Y*Y']
```

```
1  
1 0.0000
```

```
MATRIX D
```

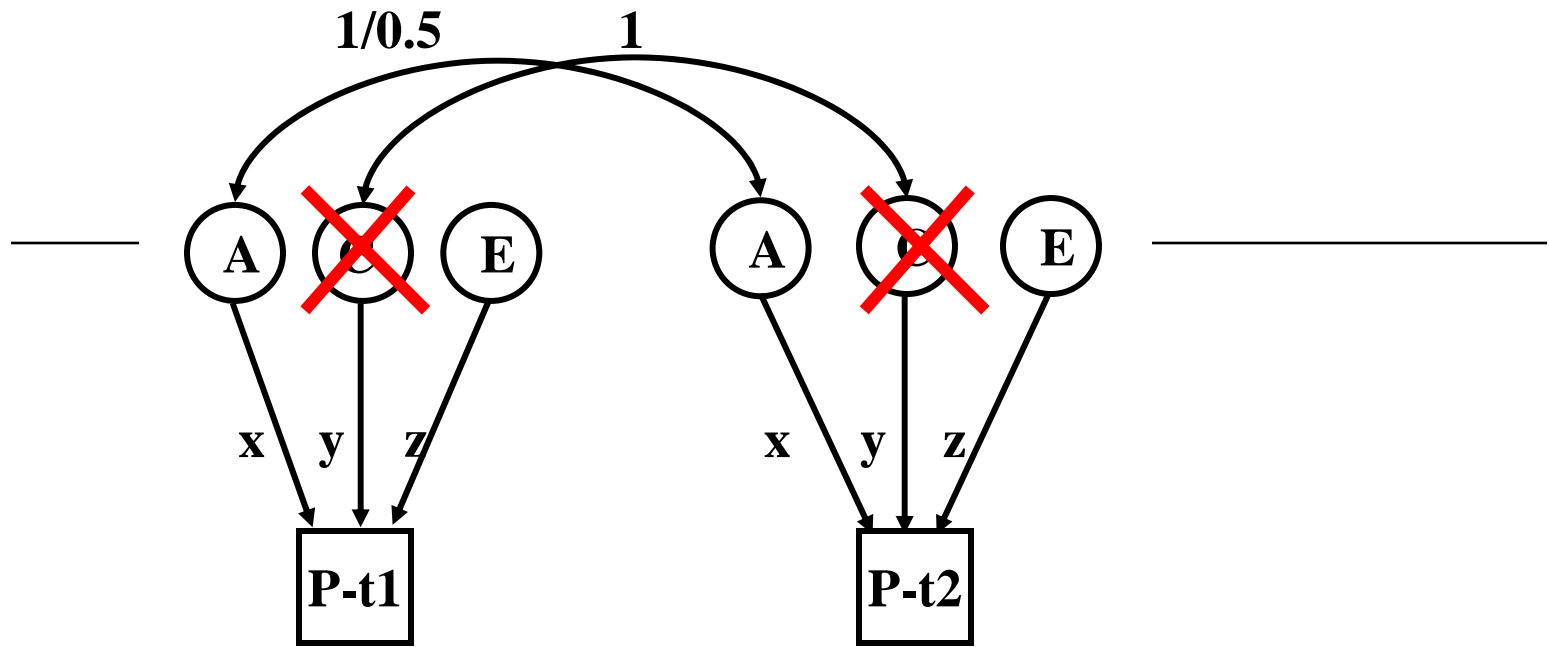
```
This is a computed FULL matrix of order 1 by 1  
[=W*W']
```

```
1  
1 0.1084
```




Testing for significance

- Drop the parameter(s) from the model or equate parameters using the multiple job option
 - Specify the matrix elements you wish to drop/equate
 - Drop A 1 1 1
 - EQ A 1 1 1 B 1 1 1
 - Compare the fit of the two models



Variance/covariance matrices

MZ

DZ

	t1	t2
t1	$a^2 + \cancel{c^2} + e^2$	$a^2 + \cancel{c^2}$
t2	$a^2 + \cancel{c^2}$	$a^2 + \cancel{c^2} + e^2$

	t1	t2
t1	$a^2 + \cancel{c^2} + e^2$	$0.5a^2 + \cancel{c^2}$
t2	$0.5a^2 + \cancel{c^2}$	$a^2 + \cancel{c^2} + e^2$



What to report

- Summary statistics
 - Usually from a simplified 'saturated' model
- Standardized estimates
 - Easier to conceptualise
 - ie 40% of the phenotypic variance vs a genetic effect of 2.84
 - Can easily be returned to original scale if summary statistics are provided



What to report

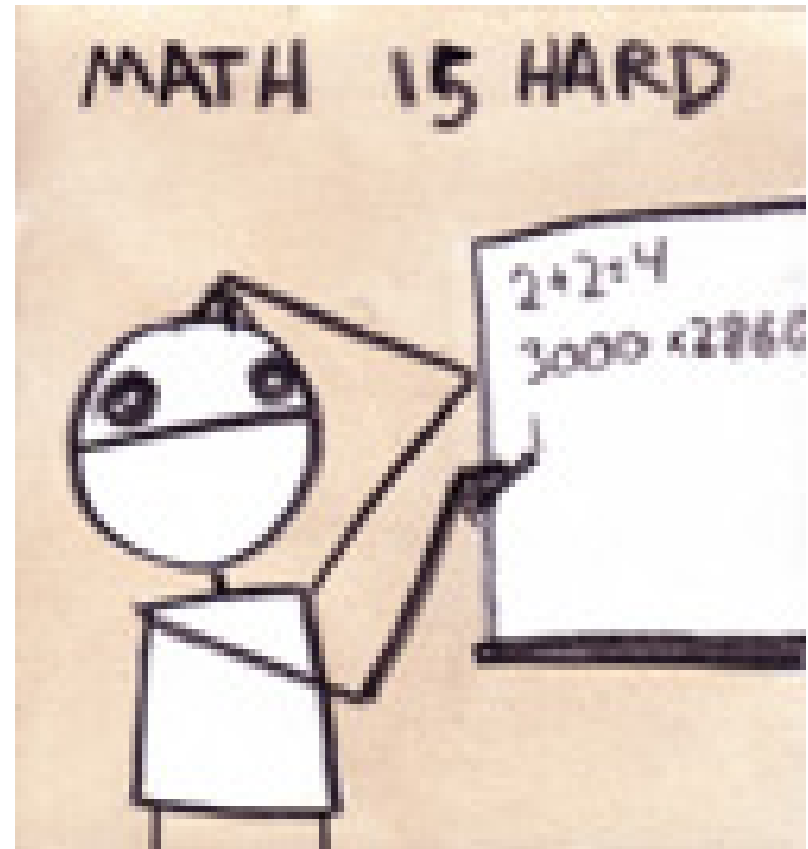
- Path coefficients
 - Very Important in multivariate analyses
 - Gives a much clearer picture of the directionality of effects
- Variance components/proportion of variance explained
- Genetic correlations



General Advice/Problem solving

- Scripting styles differ
- Check the parameter numbers
- Check the sample description
- Learn to love the manual
- Comments are your friends

Time for coffee



explodingdog.com