MKF[1]

Provisional documentation

Abstract

MKF calculates normal theory ML estimates in the time-invariant Kalman Filter (KF) using the predicition error decomposition. This is an implementation of Harvey (1996, p. 144ff). Maximalization of the loglikelihood function is carried out by NPSOL, a quasi-Newton optimization routine, using exact gradients.

The implementation includes the possibility to fit multi-group models. Here a group may consist of one (NS=1), or more cases (NS>1). Given NS>1, the cases within a group are assumed constitute an ensemble, i.e., they are assumed to be identically and independently distributed. Below we shall refer to "a group" as "a model". Parameters may be freely estimated, estimated subject to equality constraints, or fixed to known values. Standard errors of estimates are based on the central differences approximation of the Hessian using exact gradients.

Warning: Beyond a small number of applications, this program has not been subjected to very much testing. It comes with no guarantee to produce correct results, etc. etc. Please email the author, in case of detected bugs.

[1] Current version mkfm6 – dated in mkfm6.f as nov 5 2002. feb05 minor revision of manual. Bug in multigroup fixed (routine nsand). Author: Conor Dolan, c.v.dolan@uva.nl.

Time series model

MKF incorporates the following timeseries model:

$$\mathbf{y}[t] = \mathbf{S}\ \mathbf{a}[t] + \mathbf{d} + \mathbf{e}[t] + \mathbf{Z}\ \mathbf{x}[t]$$

$$\mathbf{a}[t+1] = \mathbf{H}\ \mathbf{a}[t] + \mathbf{c} + \mathbf{G}\ \mathbf{z}[t+1]$$

where t (t=1…T) indexes time.

The following are the random variables in the model:

y[t] observed random variable (NY x 1)
a[t] latent random variable (NE x 1)
e[t] latent (residual) random (NY x 1)
z[t] latent (residual) random (NE x 1)
x[t] observed fixed exogeneous regressors (NX x 1)

The time invariant parameters matrices and vector are the following (As the LISREL model bear a close resemblance to the present model, we include LISREL counterparts in parentheses, as an aid to those who are familiar with the LISREL system):

S (NY x NE matrix) regression parameters **y**[t] on **a**[t] (LISREL: Λ_y)
d (NY x 1 vector) intercept in regession of **y**[t] on **a**[t] (LISREL: τ_y)
H (NE x NE matrix) regression parameters **a**[t+1] on **a**[t] (LISREL: **B**)
c (NE x 1 vector) intercept in regression of **a**[t+1] on **a**[t] (LISREL: α)
R (NY x NY symmetric matrix) covariance matrix of **e**[t] (LISREL: Θ_ε)
Q (NE x NE symmetric matrix) covariance matrix of **z**[t] (LISREL: Ψ)
Z (NY x NX matrix) regression parameters **y**[t] on **x**[t]
G (NE x NE matrix) regression parameters **a**[t] on **z**[t]

The variables **z**[t] and **e**[t] are assumed to be (multi-) normally distributed, or (multi-) normally distributed, conditional on **x**[t], if present. The multi-normal state vector at t=0, **a**[0], and its covariance matrix, **P**[0], are assumed to be known, and has to be provided by the user.

The components of these matrices and vectors may include fixed parameters, and parameters to-be-estimated, i.e., free parameters. Free parameters may be subject to equality constraints.

Multi-subject model (NS>1)

We employ the following simple generalization to multiple subjects. Let i (i=1…N) be the index of case. The model in this case is simply:

$$\mathbf{y}[t_i]_i = \mathbf{S}\ \mathbf{a}[t_i]_i + \mathbf{d} + \mathbf{e}[t_i]_i + \mathbf{Z}\ \mathbf{x}[t_i]_i$$

$$\mathbf{a}[\{t+1\}_i]_i = \mathbf{H}\ \mathbf{a}[t_i]_i + \mathbf{c} + \mathbf{G}\ \mathbf{z}[\{t+1\}_i]_i$$

The time index has a subject index to communicate the fact that each subject or case need not have the same length of time series. Note that the parameter matrices do not have time or subject indices. These are assumed to be constant over time and over the NS subjects. So in the multi-subject case, we specify a single model for a sample of NS cases. These NS cases are presented by NS timeseries (not necessarily of equal length). We call such a collection of independently and identically distributed timeseries an ensemble.

Multi-subject (NS>1), multi-model (NM>1)

We employ the following generalization of the time series model to multiple groups. Let i (i=1…NS) again index subject or case and let j index the model (j=1…NM). The model in this case is:

$$\mathbf{y}[t_{ij}]_{ij} = \mathbf{S}_j\ \mathbf{a}[t_{ij}]_{ij} + \mathbf{d}_j + \mathbf{e}[t_{ij}]_{ij} + \mathbf{Z}_j\ \mathbf{x}[t]_{ij}$$

$$\mathbf{a}[\{t+1\}_{ij}]_{ij} = \mathbf{H}_j\ \mathbf{a}[t_{ij}]_{ij} + \mathbf{c}_j + \mathbf{G}_j\ \mathbf{z}[\{t+1\}_{ij}]_{ij}$$

In addition:
$\mathbf{R}_j$ covariance matrix of $\mathbf{e}[t]_{.j}$ in model j
$\mathbf{Q}_j$ covariance matrix of $\mathbf{z}[t]_{.j}$ in model j

The dimensions of the parameter matrices / vectors $\mathbf{S}_j$, $\mathbf{d}_j$, $\mathbf{H}_j$, $\mathbf{c}_j$, $\mathbf{Z}_j$, $\mathbf{G}_j$, $\mathbf{R}_j$, and $\mathbf{Q}_j$ may vary over models j=1…NM, i.e. the dimensions of the times series as defined by NY, NE, NX may vary over the models.

The elements in the parameter matrices / vectors $\mathbf{S}_j$, $\mathbf{d}_j$, $\mathbf{H}_j$, $\mathbf{c}_j$, $\mathbf{R}_j$, $\mathbf{Z}_j$, $\mathbf{Q}_j$, $\mathbf{G}_j$, and $\mathbf{P}_j$ may vary over models, or may be constrained to be equal over models (j=1…NM).

See Table below for examples of possible analyses.

Table: Examples of some possibilities.

	NS	NM
1)single subject	1	1
2)10 subject in an ensemble	10	1
3)two subject not (necessarily*) in an ensemble:	1 (in each NM)	2
4) 5 subjects in two distinct ensembles	5 (in each NM)	2
5) 5 subject not (necessarily) in an ensemble	1 (in each NM)	5

*Whether the subjects do form an ensemble may be investigated. E.g., appropriate likelihood ratio tests may indicate that the subjects do not differ in the values of their model parameters.

So one may consider a collection of time-series as realizations of a single process, i.e., poolable in the sense that they form an ensemble. One may consider, say, two timeseries as distinct, not an ensemble (NM=2). The possibility of analyzing these simultaneously allows one to investigate possible parameter equalities over the timeseries. This can be done by fitting the model with and without relevant equality constraints and conducting a likelihood ratio test. An ensemble may consist of a number of time series, but may a single time series may be viewed as an ensemble of 1.

Estimation

MKF calculates normal theory ML estimates in the time-invariant Kalman Filter (KF) using the predicition error decomposition. This present implementation is based on Harvey (1996, p. 144ff). Ellen Hamaker has provided additional documentation, which includes an explanation of this method, as well as useful notes on model specification of dynamic factor models.

Maximalization of the loglikelihood function is carried out by NPSOL, a quasi-Newton optimization routine, using exact gradients. Harvey (1996) provides recursive expressions for these. Standard errors are calculated by means of central finite difference approximation of the observed Information matrix.

Missingness

Some or all components of the vector **y**[t] may be missing. The missing code is given in the input file by **mi=*** (e.g., mi=-999). Fixed regressors, if present, may **not** include missing data.

Running the program

To run the program, one has to prepare an input file, which specifies the model, and a data file, which contains the data. Let's call in the input file *infile_1* and the output file *outile_1*. The program is run from a so-called DOS window (under WINDOWS) by typing, at the DOS prompt:

```
mkfm6 < infile_1 > outfile_1
```

The input file should be prepared as specified below. The output file will contain the results of the analysis, or errors, if encountered. Error message will hopefully provide some useful hint of what went wrong.

NOTE: Certain parameter maxima (e.g., number of subjects, number of observed variables, length of timeseries) may be reached. If so please contact the author for a adapted version.

Input requirements

NB: capitalization as shown should be adhered to. For instance the program expect "title" not Title or TITLE.

NOTE: Certain parameter maxima (e.g., number of subjects, number of observed variables, length of timeseries) may be reached. If so please contact the author for an adapted version.

Known bugs (number=1):

1) Add RETURN to last line of input file. Last line should thus be an empty line. Same may apply to the data file.

I Outline of input file for single model (or group).

text in input file	annotation
title	*title*
nm=1 se=yes it=100	*line number of models, st. error option*
!	
mo=1 ny=* ne=* nx=*	*define model*
!	
df=* rf=* ns=* mi=*	*define data + files*
!	
S=* R=* H=* Q=* d=* c=* G=* Z=* P=1	*indicators of parameters*
!	
{ define matrices }	*provide model matrices*
!	
st	*starting values*
{starting values}	
ub	*upper bounds*
{upper bounds}	
lb	*lower bounds*
{lower bounds}	

Explanation

title is the title (has to be present)

nm=1 number of models (equals one here, but not generally)

se=yes calculate standard errors (se=no, do not calculate standard errors)

! lines starting with ! are skipped

it=100 number of iteration (if absent, default is **it=1000**).

mo=1 model number 1 (i.e, model **1** of **nm**, here)

ny= number of observed variables in **y**

ne= length of state vector, number of latent variables **a**

nx= number of fixed regressors **x**

!

df= data input file containing time series **y**[t] (max 12 characters)

rf= data output file contains the estimated states **a**[t] (max 12 characters)

! note: if **rf=no** states are not saved to file.

ns= number of subjects in model 1 (**ns=1** or **ns>0**)

mi= missing value indicator (e.g. mi=-999.)

!

S=1 means matrix **S** is used in the model

S=0 means matrix **S** is not used in the model

* Same applies to **R**, **H**, **Q**, **d**, **c, Z**

NB: The matrix **P** , the covariance matrix of **a**[0], has to be present (**P=0** generates an error).

If **S=1**, then in the section "define matrices", the fixed and free matrices **S** are specified, as follows:

S fi

{ny x ne matrix – fixed elements, fortran type real}

S fr

{ny x ne matrix – free elements, fortran type integer}

or if S is diagonal:

S fi di

{diagonal of ny x ny matrix – fixed elements, fortran type real diagonals}

S fr di

{diagonal of ny x ny matrix – free elements, fortran type integer diagonals}

idem **R** (ny x ny lower triangular)
idem **H** (ne x ne full)
idem **Q** (ne x ne lower triangular)
idem **d** (1 x ny vector)
idem **c** (1 x ne vector)
idem **Z** (ny x nx full)
idem **G** (ne x nx full)
idem **P** (ne x ne lower triagular)

In the matrices following **S fr** (**R fr** etc. etc.), integer indicate the parameter to estimated. For instance:

S fi

```
0 0
0 0
0 0
S fr
2 4
6 8
10 12
```

indicates that 6 distinct parameters are to be estimated in S. The following specification includes fixed parameters in positions 1 1 and 3 2. The parameters in these positions are fixed to equal 1.

```
S fi
1 0
0 0
0 1
S fr
0 4
6 8
10 0
```

Equality constraints are specified by using the same integer:

```
S fi
1 0
0 0
0 1
S fr
0 4
6 4
6 0
```

We would now estimate only 2 parameters (1 in position 2 1 and 3 1, and 1 in positions 1 2 and 2 2). The 1's in positions 1 1 and 3 2 are fixed.

Starting values are specified at the end of the file. The actual starting values follow the keyword **st.** The starting values should correspond in order to the integer used to indicate the position of the free parameters. For instance, given:

```
S fi
1 0
0 0
0 1
S fr
0 4
6 8
10 0
```

st
 1 .2 .5 1

would start element 1 2 of S at 1, element 2 2 of S at .5, etc.

Lower and upper bounds follow the **lb** and **ub** keywords. These have to be present. If bounds are not required the upper and lower bounds should be chosen librally (e.g., -1000 +1000)

Note 1: The matrix **P** represents the fixed (known) covariance matrix of **a**[0], **P**[0]. **P=1** should always be specified. **P fi** has to be present. **P fr** has to be followed by (ne x ne) lower triangular of zeros. Any deviation from this will generate an error and terminate the program (or given **P fi di** and **P fr di**, vectors).

Note 2: The possibility of specifying a matrix as diagonal is applicable to each model matrix. Generally this will be used in the case of symmetric matrices (**R**, **Q**, **P**). It is not applicable to the vectors **d** and **c**, which, as vectors, cannot be diagonal.

II Outline of input file for <u>multiple models (or groups)</u>, say nm=3.

text in input file   annotation

```
title                                 title
nm=3 se=yes                           line number of models
!
! first model ---------------------------------------
!
mo=1 nx=* ny=* ne=*                   define model
!
df=* rf=* ns=* mi=*                   define data + files
!
S=* R=* H=* Q=* d=* c=* Z=* G=* P=1   indicate parameters
!
{ define matrices }                   provide model matrices
!
! next model ---------------------------------------
!
mo=2 ny=* ne=* nx=*                   define model
!
df=* rf=* ns=* mi=*                   define data + files
!
S=* R=* H=* Q=* d=* c=* G=1 Z=* P=1   indicate parameters
!
{ define matrices }                   provide model matrices
!
! final model ---------------------------------------
!
mo=3 ny=* ne=* nx=*                   define model
!
df=* rf=* ns=* mi=*                   define data + files
!
S=* R=* H=* Q=* d=* c=* G=1 Z=* P=1   indicate parameters
!
{ define matrices }                   provide model matrices
!
!
! -----------------------------------------------------
st                                    starting values
{starting values}
ub                                    upper bounds
{upper bounds}
lb                                    lower bounds
{lower bounds}
```

Data file specified in df=, nx=0.

Suppose the file containing the observed timeseries is called tsdat1 (**df=tsdat1**). Data is read in free format. This file should have the following structure, given **nx**=0 (no fixed regressors):

nt	number of timeseries
a[0]	ne row vector a[0]
y[1]	ny row vector y at t=1
y[2]	
y[3]	
...	
y[nt]	ny row vector y at t=nt.

In the analyses of an ensemble of timeseries, this structure should be repeated for each case. E.g. **ns=2**:

nt	number of timeseries case 1
a[0]	ne row vector a[0] case 1
y[1]	ny row vector y at t=1 case 1
y[2]	
y[3]	
...	
y[nt]	ny row vector y at t=nt case 1.
nt	number of timeseries case 2
a[0]	ne row vector a[0] case 2
y[1]	ny row vector y at t=1 case 2
y[2]	
y[3]	
...	
y[nt]	ny row vector y at t=nt case 2.

Data file specified in df=, nx>0.

Suppose the file containing the observed timeseries is called tsdat1 (**df=tsdat1**). Data is read in free format. This file should have the following structure, given **nx**=0 (no fixed regressors):

nt	number of timeseries
a[0]	ne row vector a[0]
y[1] **x**[1]	ny row vector y & nx row vector at t=1
y[2] **x**[2]	
y[3] **x**[3]	
...	
y[nt] **x**[nt]	ny row vector y & nx row vector at t=nt.

In the analyses of an ensemble of timeseries, this structure should be repeated for each case. E.g. **ns=2**:

nt	number of timeseries case 1
a[0]	ne row vector a[0] case 1
y[1] **x**[1]	ny row vector y and nx xat t=1 case 1
y[2] **x**[2]	
y[3] **x**[3]	
...	
y[nt] **x**[nt]	ny row vector y and nx x at t=nt case 1.
nt	number of timeseries case 2
a[0]	ne row vector a[0] case 2
y[1] **x**[1]	ny row vector y and nx x at t=1 case 2
y[2] **x**[2]	
y[3] **x**[3]	
...	
y[nt] **x**[nt]	ny row vector y and nx x at t=nt case 2.

Example 1: MA(1) – specification 1 $y_t = \mu + \eta_t + \theta\ \eta_{t-1}$

```
title example  simulated

nm=1 se=yes
mo=1 ny=1 ne=2 nx=0
df=ma1.dat rf=no ns=1 mi=-999.
S=1 R=0 H=1 Q=1 d=1 c=0 P=1 Z=0 G=1

d fi
0
d fr
3

H fi
0 0
0 0
H fr
0 0
1 0

Q fi di
0 0
Q fr di
2 0

G fi di
1 1
G fr di
0 0

P fi di
2 2
P fr di
0 0

S fi
1 1
S fr
0 0

st
  .3  1  0
lb
 -1 .001    -1
ub
  1 10 10
```

Example 1: MA(1) – specification 2 .

```
title example simulated

nm=1 se=yes
mo=1 ny=1 ne=2 nx=0
df=ma1.dat rf=no ns=1 mi=-999.
S=1 R=0 H=1 Q=1 d=1 c=0 P=1 Z=0 G=1

d fi
0
d fr
3

H fi
0 0
1 0
H fr
0 0
0 0
```

```
Q fi di
0 0
Q fr di
2 0

G fi di
1 1
G fr di
0 0

P fi di
2 2
P fr di
0 0

S fi
1 1
S fr
0 1

st
  .3  1  0
lb
 -1 .001    -1
ub
  1 10 10
```

Example 2: MA(2) – specification 1.

$$y_t = \mu + \eta_t + \theta_1 \eta_{t-1} + \theta_2 \eta_{t-2}$$

```
title example simulated

nm=1 se=yes
mo=1 ny=1 ne=3 nx=0
df=ma2.dat rf=no ns=1 mi=-999.
S=1 R=0 H=1 Q=1 d=1 c=0 P=1 Z=0 G=1

d fi
0
d fr
40

H fi
0 0 0
0 0 0
0 0 0
H fr
0 0 0
10 0 0
0 20 0

Q fi di
0 0 0
Q fr di
30 0 0

G fi di
1 1 1
G fr di
0 0 0

P fi di
5 5 5
P fr di
0 0 0

S fi
1 1 1
S fr
0 0 0
```

```
st
  -.27 .17 1  0
lb
 -.8 -.8 .001    -1
ub
  .8 .8 2 1
```

Example 2: MA(2) – specification 2.

```
title example simulated

nm=1 se=yes
mo=1 ny=1 ne=3 nx=0
df=ma2.dat rf=no ns=1 mi=-999.
S=1 R=0 H=1 Q=1 d=1 c=0 P=1 Z=0 G=1

d fi
0
d fr
4

H fi
0 0 0
1 0 0
0 1 0
H fr
0 0 0
0 0 0
0 0 0

Q fi di
0 0 0
Q fr di
3 0 0

G fi di
1 1 1
G fr di
0 0 0


P fi di
5 5 5
P fr di
0 0 0

S fi
1 1 1
S fr
0 1 2

st
  -.31 .2 1  0
lb
 -.8 .1 .001    -1
sub
  .8 .8 2 1
```

Note MA(2) specification 2 gives the following results (true values θ_1=-.3, θ_2 =.2):

```
S parameters
   1.000  -0.276   0.175

H parameters
   0.000   0.000   0.000
   1.000   0.000   0.000
   0.000   1.000   0.000

Q parameters - diagonal
   1.062   0.000   0.000
```

But MA(2) specification 1 given these results (true values θ_1=-.3, θ_2 =.2):

S parameters

1.000	1.000	1.000

H parameters

0.000	0.000	0.000
-0.280	0.000	0.000
0.000	-0.612	0.000

Q parameters - diagonal

1.063	0.000	0.000

The two are equivalent. -.280 is the estimate of θ_1 (true value -.30) and -.612 is the estimate of θ_2/θ_1 = .2/-.3 = -.667. The fact that θ_2/θ_1 is estimated rather than θ_2 is due to the nature of the model specification.

Example 3: AR(2) – specification.

$$\eta_t = \mu + \phi_1 \eta_{t-1} + \phi_2 \eta_{t-2} + \zeta_t$$

```
title example simulated

nm=1 se=yes
mo=1 ny=1 ne=2 nx=0
df=ar2.dat rf=no ns=1 mi=-999.
S=1 R=0 H=1 Q=1 d=1 c=0 P=1 Z=0 G=1

G fi di
1 1
G fr di
0 0

d fi
0
d fr
4

H fi
0 0
1 0
H fr
1 2
0 0

Q fi di
1 0
Q fr di
3 0

P fi di
5 5
P fr di
0 0

S fi
1 0
S fr
0 0

st
  .1 .1 1  0
lb
 -.8 -.8 .001    -1
ub
  .8 .8 2 1
```

Example 4: ARMA 11 – specification.

$$\eta_t = \mu + \phi\, \eta_{t-1} + \theta\, \zeta_{t-1} + \zeta_t$$

```
title example simulated

nm=1 se=yes
mo=1 ny=1 ne=4 nx=0
df=arma11.dat rf=no ns=1 mi=-999.
S=1 R=0 H=1 Q=1 d=1 c=0 P=1 Z=0 G=1

d fi
0
d fr
4
H fi
0 0 0 0
1 0 0 0
0 0 0 0
0 0 1 0
H fr
1 0 2 0
0 0 0 0
0 0 0 0
0 0 0 0
G fi
1 0 0 0
0 0 0 0
1 0 0 0
0 0 0 0
G fr
0 0 0 0
0 0 0 0
0 0 0 0
0 0 0 0
Q fi di
0 0 0 0
Q fr di
3 0 0 0
P fi di
5 5 5 5
P fr di
0 0 0 0
S fi
1 0 0 0
S fr
0 0 0 0

st
  .3 .3 1  0
lb
 -.8 -.8 .001    -1
ub
.8 .8 2 1
```

Example 5: Dynamic Factor Model (Generalized P-technique) – specification. (Molenaar, 1885, Psychometrika).

$$y_t = \mu + \Lambda_0 \zeta_t + \Lambda_1 \zeta_{t-1} + \varepsilon_t$$

```
title example  simulated

nm=1 se=yes
mo=1 ny=4 ne=2 nx=0
df=datdfb rf=no ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=0 P=1 Z=0 G=1

R fi di
0 0 0 0
R fr di
10 11 12 13

H fi
0 0
1 0
H fr
0 0
0 0

G fi di
1 1
G fr di
0 0

Q fi
1
0 0
Q fr
0
0 0

P fi
10
0 10
P fr
0
0 0

d fi
0 0 0 0
d fr
20 21 22 23

S fi
0 0
0 0
0 0
0 0
S fr
1 5
2 6
3 7
4 8
st
 .6 .7 .6 .7    .6 .4 .6 .4  .3 .3 .3 .3  .01 .01 .01 .01
lb
.001 .001 .001 .001  .001 .001 .001 .001  .001 .001 .001 .001  -10 -10 -10 -10
ub
2 2 2 2 2 2 2 2 2 2 2 2 10 10 10 10
```

Example 6: Multiple indicator ar(1) – specification.

```
title example  simulated

nm=1 se=yes nr=5
mo=1 ny=3 ne=1 nx=0
df=mdat2 rf=no ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=1 P=1 Z=0 G=1

G fi di
1
G fr di
0

R fi di
 0.8868 0.8868 0.8868
R fr di
15 16  17

H fi
.6437
H fr
30

Q fi di
.5604
Q fr di
35

P fi
10
P fr
0

c fi
5
c fr
5

d fi
5 5 5
d fr
5 5 5

S fi
1
.9788
.9788
S fr
0
10
12

st
 1  1 1 1 1 1 .6 .4
lb
 -5  -2 -2 -2 -2  .1 -.9 .1
ub
  5  2 2 2 2 10  .9 3
```

Example 7: Multiple indicator ar(1) in two subjects– specification. Note different values of the parameters d are estimated in each subject (if all parameters were equal, then it would be easier to specify nm=1 and ns=2).

```
title example  simulated

nm=2 se=yes
mo=1 ny=3 ne=1 nx=0
df=mdat2 rf=tmp ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=1 P=1 G=1 Z=0

G fr
0
G fi
```

```
1

c fr
0
c fi
0

R fi di
 0 0 0
R fr
15 20 25

H fi
.6437
H fr
30

Q fi
.5604
Q fr
35

P fi
100
P fr
0

d fi
0 0 0
d fr
6 6 6

S fi
1
.9788
.9788
S fr
0
10
15

mo=2 ny=3 ne=1 nx=0
df=mdat1 rf=tmp ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=1 P=1 G=1 Z=0

G fr
0
G fi
1

c fr
0
c fi
0

R fi
 0.8868 0 0.8868 0 0 0.8868
R fr
15  0 20  0 0 25

H fi
.6437
H fr
30

Q fi
.5604
Q fr
35

P fi
100
P fr
0

d fi
0 0 0
```

```
d fr
5 5 5

S fi
1
.9788
.9788
S fr
0
10
15

st
 1 1  2 2    1 1 1 .3 .5
lb
 0  0  -2 -2 .1 .1 .1 -.9 .1
ub
 120 100  3 3 10 10 10 .9 3
```

Example 8: Dynamic Factor Model (Generalized P-technique) – specification. No correlated residuals.

```
title simulatiedata
nm=1 se=yes
mo=1 ny=5 ne=7 nx=0
df=xsdat rf=states ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=0 P=1 Z=0 G=1

G fi di
1 1 1 1 1 1 1
G fr di
0 0 0 0 0 0 0

S fi
0 0 1 0 0 0 0
0 0 0 1 0 0 0
0 0 0 0 1 0 0
0 0 0 0 0 1 0
0 0 0 0 0 0 1
S fr
2 7  0 0 0 0 0
3 8  0 0 0 0 0
4 9  0 0 0 0 0
5 10  0 0 0 0 0
6 11  0 0 0 0 0

Q fi di
1 0  0 0 0 0 0
Q fr di
0 0 17 18 19 20 21

H fi fu
0 0  0 0 0 0 0
1 0  0 0 0 0 0
0 0  0 0 0 0 0
0 0  0 0 0 0 0
0 0  0 0 0 0 0
0 0  0 0 0 0 0

0 0  0 0 0 0 0
H fr fu
0 0  0 0 0 0 0
0 0  0 0 0 0 0
0 0  0 0 0 0 0
0 0  0 0 0 0 0
0 0  0 0 0 0 0
0 0  0 0 0 0 0
0 0  0 0 0 0 0

d fi
0 0 0 0 0
d fr
12 13 14 15 16
```

```
R fi di
0 0 0 0 0
R fr di
0 0 0 0 0

P fi sy
100
0 100
0 0 0
0 0 0 0
0 0 0 0 0
0 0 0 0 0 0
0 0 0 0 0 0 0
P fr sy
0
0 0
0 0 0
0 0 0 0
0 0 0 0 0
0 0 0 0 0 0
0 0 0 0 0 0 0
st
1 1 1 1 1
1 1 1 1 1
0 0 0 0 0
1 1 1 1 1
lb
-100 -100 -100 -100 -100
-100 -100 -100 -100 -100
-100 -100 -100 -100 -100
.0001 .0001 .0001 .0001 .0001
ub
1000 1000 1000 1000 1000
1000 1000 1000 1000 1000
1000 1000 1000 1000 1000
1000 1000 1000 1000 1000
```

Example 8b: Dynamic Factor Model (Generalized P-technique) – specification. With correlated residuals (as discussed in Molenaar 1985, Psychometrika).

```
title simulatiedata
nm=1 se=yes
mo=1 ny=5 ne=7 nx=0
df=xsdat rf=states ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=0 P=1 Z=0 G=1

G fr di
0 0 0 0 0 0 0
G fi di
1 1 1 1 1 1 1

S fi
0 0 1 0 0 0 0
0 0 0 1 0 0 0
0 0 0 0 1 0 0
0 0 0 0 0 1 0
0 0 0 0 0 0 1
S fr
2 7  0 0 0 0 0
3 8  0 0 0 0 0
4 9  0 0 0 0 0
5 10  0 0 0 0 0
6 11  0 0 0 0 0

Q fi di
1 0  0 0 0 0 0
Q fr di
0 0 17 18 19 20 21

H fi fu
0 0  0 0 0 0 0
1 0  0 0 0 0 0
```

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

0 0  0 0 0 0 0
H fr fu
0 0  0 0 0 0 0
0 0  0 0 0 0 0
0 0 31 0 0 0 0
0 0  0 32 0 0 0
0 0  0 0 33 0 0
0 0  0 0 0 34 0
0 0  0 0 0 0 35

d fi
0 0 0 0 0
d fr
12 13 14 15 16


R fi di
0 0 0 0 0
R fr di
0 0 0 0 0

P fi sy
100
0 100
0 0 0
0 0 0 0
0 0 0 0 0
0 0 0 0 0 0
0 0 0 0 0 0 0
P fr sy
0
0 0
0 0 0
0 0 0 0
0 0 0 0 0
0 0 0 0 0 0
0 0 0 0 0 0 0

st
1 1 1 1 1
1 1 1 1 1
0 0 0 0 0
1 1 1 1 1
.1 .1 .1 .1 .1
lb
-100 -100 -100 -100 -100
-100 -100 -100 -100 -100
-100 -100 -100 -100 -100
.0001 .0001 .0001 .0001 .0001
-100 -100 -100 -100 -100
ub
1000 1000 1000 1000 1000
1000 1000 1000 1000 1000
1000 1000 1000 1000 1000
1000 1000 1000 1000 1000
1000 1000 1000 1000 1000
```

Example 9a: Multiple indicator AR(1) model – specification, with regression of indicators on 2 fixed regressors. The first regressor is the unit vector, the second is a linear constrast: 1,2,3,4...

Top of the file xdatr looks like this:

```
100
0
```

```
  1.0667770e+001  9.7491203e+000  1.0681997e+001  1.0000044e+001  1.0000000e+000
1.0000000e+000
  1.1686892e+001  9.3924026e+000  9.9047913e+000  1.0109826e+001  1.0000000e+000
2.0000000e+000
  9.8631308e+000  1.0418394e+001  8.5172602e+000  9.9299308e+000  1.0000000e+000
3.0000000e+000
  8.9038781e+000  9.1235257e+000  1.0232686e+001  1.1151494e+001  1.0000000e+000
4.0000000e+000
  8.9372764e+000  8.7090719e+000  8.7399348e+000  9.1584549e+000  1.0000000e+000
5.0000000e+000
  8.6763339e+000  9.7141843e+000  9.3876403e+000  9.2216587e+000  1.0000000e+000
6.0000000e+000
  9.7483628e+000  9.4155161e+000  9.2537202e+000  8.6030098e+000  1.0000000e+000
7.0000000e+000
  1.0367206e+001  8.6947878e+000  9.7328273e+000  8.4316556e+000  1.0000000e+000
8.0000000e+000
  1.0631671e+001  1.0734796e+001  8.9309345e+000  8.8141668e+000  1.0000000e+000
9.0000000e+000

.....


  1.7711814e+001  2.0028421e+001  1.7619503e+001  1.8463804e+001  1.0000000e+000
1.9600000e+002
  2.1033009e+001  2.0433182e+001  2.1585509e+001  2.1141994e+001  1.0000000e+000
1.9700000e+002
  2.1888676e+001  2.1194893e+001  2.2069753e+001  2.2375154e+001  1.0000000e+000
1.9800000e+002
  2.0037252e+001  2.2353028e+001  2.0611591e+001  2.0975476e+001  1.0000000e+000
1.9900000e+002
  2.1064460e+001  2.1007033e+001  2.1350417e+001  2.1443359e+001  1.0000000e+000
2.0000000e+002

title example  simulated

nm=1 se=yes
mo=1 ny=4 ne=1 nx=2
df=xdatr rf=no ns=1 mi=-999.
S=1 R=1 H=1 Q=1 d=1 c=0 Z=1 P=1 G=1

R fi
 0 0 0 0 0 0 0 0 0 0
R fr
10
0 15
0 0 20
0 0 0 25

H fi
.6
H fr
30

G fi
1
G fr
0
Q fi
0
Q fr
35

P fi
100
P fr
0

Z fi
 0 0.05
 0 0.05
 0 0.05
 0 0.05
Z fr
61 65
62 66
63 67
64 68
```

```
d fi
0 0 0 0
d fr
0 0 0 0

S fi
1
0
0
0
S fr
0
40
45
50

st
  1 1 1 1   .5 .5   1 1 1  10 10 10 10 .1 .1 .1 .1
lb
 .1 .1 .1 .1 -1 .1   .1 .1 .1   -1 -1 -1 -1 -1 -1 -1 -1  -1 -1 -1 -1
ub
 6 6 6 6   6 6    6 6 6   20 20 20 20 20 20 20 20 20 20 20 20
```

Example 9b: Multiple indicator AR(1) model – specification, with regression of indicators on 2 fixed regressors. Regression on the fixed unit vector is fixed to zero. Instead, we use the parameter vector d to estimate the intercepts. With model is equivalent to the model in Example 9a.

```
title example  simulated

nm=1 se=yes
mo=1 ny=4 ne=1 nx=2
df=xdatr rf=no ns=1 mi=-999.
S=1 R=1 H=1 Q=1 d=1 c=0 Z=1 P=1 G=1

R fi
 0 0 0 0 0 0 0 0 0 0
R fr
10
0 15
0 0 20
0 0 0 25

H fi
.6
H fr
30

G fi
1
G fr
0
Q fi
0
Q fr
35

P fi
100
P fr
0

Z fi
 0 0.05
 0 0.05
 0 0.05
 0 0.05
Z fr
 0 65
 0 66
```

```
 0 67
 0 68

d fi
0 0 0 0
d fr
61 62 63 64

S fi
1
0
0
0
S fr
0
40
45
50

st
  1 1 1 1   .5 .5   1 1 1  10 10 10 10 .1 .1 .1 .1
lb
 .1 .1 .1 .1 -1 .1   .1 .1 .1   -1 -1 -1 -1 -1 -1 -1 -1  -1 -1 -1 -1
ub
 6 6 6 6   6 6    6 6 6   20 20 20 20 20 20 20 20 20 20 20 20
```

OUTPUT 9a

```
max nm= 3 nt=5000 ns=  10 ny=30 nx= 5 ne=30 npar=400



Read from input file

title example  simulated
nm=1 se=yes

mo=1 ny=4 ne=1 nx=2
df=xdatr rf=no ns=1 mi=-999.
S=1 R=1 H=1 Q=1 d=1 c=0 Z=1 P=1 G=1

                          ===================
                          MKFv1     Nov. 2002
                          ===================

title example  simulated

Model  1 of  1

S fr parameters (nonzero)
   0
   7
   8
   9

R fr parameters (nonzero)
   1
   0   2
   0   0   3
   0   0   0   4

H fr parameters (nonzero)
   5

Q fr parameters (nonzero)
   6

Z fr parameters (nonzero)
  10  14
  11  15
  12  16
  13  17

DATA SUMMARY
```

```
MODEL  1 of  1 NY= 4 NX= 2 NE= 1 Ncases=   1 START=   1 END=   1

CASE    1    T= 100 N of T missing=   0

   State_0    0.00
   var          1         2         3         4
   %miss   0.0000   0.0000   0.0000   0.0000
   mean     12.29    12.19    12.29    12.26
   var       3.32     3.51     3.43     3.36
   std       1.82     1.87     1.85     1.83
   min       8.45     7.62     8.32     8.13
   max      15.59    16.90    16.89    15.75

Number of fixed regressors  2



 N of penalized error - during npsol
 --- Density less/equal zero     20 times


 ML parameter estimates
 nr   1            0.86126 g     0.000003 se       0.1389 t       6.20
 nr   2            0.88369 g    -0.000003 se       0.1459 t       6.06
 nr   3            1.10187 g     0.000003 se       0.1836 t       6.00
 nr   4            0.83124 g     0.000004 se       0.1482 t       5.61
 nr   5            0.76914 g    -0.000002 se       0.0820 t       9.38
 nr   6            0.24859 g    -0.000004 se       0.0926 t       2.68
 nr   7            1.00515 g    -0.000003 se       0.1878 t       5.35
 nr   8            1.15976 g     0.000005 se       0.2115 t       5.48
 nr   9            1.14253 g    -0.000006 se       0.1967 t       5.81
 nr  10            9.89560 g    -0.000009 se       0.5076 t      19.50
 nr  11            9.70721 g     0.000001 se       0.5107 t      19.01
 nr  12           10.14058 g     0.000003 se       0.5867 t      17.28
 nr  13            9.94585 g     0.000002 se       0.5695 t      17.46
 nr  14            0.04708 g    -0.000419 se       0.0083 t       5.65
 nr  15            0.04881 g     0.000075 se       0.0084 t       5.82
 nr  16            0.04215 g     0.000164 se       0.0096 t       4.37
 nr  17            0.04530 g     0.000152 se       0.0093 t       4.85

 Logl        -324.192 -2xLogL         648.385 Inform(NPSOL)  0

title example  simulated

Model  1 of  1

S parameters
    1.000
    1.005
    1.160
    1.143

R parameters
    0.861
    0.000    0.884
    0.000    0.000    1.102
    0.000    0.000    0.000    0.831

H parameters
    0.769

Q parameters
    0.249

P parameters
  100.000

d parameters
    0.000    0.000    0.000    0.000

Z parameters
    9.896    0.047
    9.707    0.049
   10.141    0.042
    9.946    0.045

G parameters
```

```
    1.000

P(t|t) error cov
    0.122


start: 16 48 11 end: 16 48 17 date: 30  6 2003

start getse: 16 48 15 end getse: 16 48 17
```

output 9b

```
max nm= 3 nt=5000 ns=  10 ny=30 nx= 5 ne=30 npar=400



Read from input file

title example  simulated
nm=1 se=yes

mo=1 ny=4 ne=1 nx=2
df=xdatr rf=no ns=1 mi=-999.
S=1 R=1 H=1 Q=1 d=1 c=0 Z=1 P=1 G=1

                          ===================
                          MKFv1     Nov. 2002
                          ===================

title example  simulated

Model  1 of  1

S fr parameters (nonzero)
   0
   7
   8
   9

R fr parameters (nonzero)
   1
   0   2
   0   0   3
   0   0   0   4

H fr parameters (nonzero)
   5

Q fr parameters (nonzero)
   6

d fr parameters (nonzero)
  10  11  12  13

Z fr parameters (nonzero)
   0  14
   0  15
   0  16
   0  17

DATA SUMMARY

MODEL  1 of  1 NY= 4 NX= 2 NE= 1 Ncases=   1 START=   1 END=   1

CASE    1    T= 100 N of T missing=   0

   State_0    0.00
   var          1        2        3        4
   %miss   0.0000   0.0000   0.0000   0.0000
   mean     12.29    12.19    12.29    12.26
   var       3.32     3.51     3.43     3.36
   std       1.82     1.87     1.85     1.83
   min       8.45     7.62     8.32     8.13
   max      15.59    16.90    16.89    15.75
```

```
Number of fixed regressors  2


 ML parameter estimates
 nr   1            0.86126 g     0.000003 se       0.1389 t       6.20
 nr   2            0.88369 g    -0.000003 se       0.1459 t       6.06
 nr   3            1.10187 g     0.000003 se       0.1836 t       6.00
 nr   4            0.83124 g     0.000004 se       0.1482 t       5.61
 nr   5            0.76914 g    -0.000002 se       0.0820 t       9.38
 nr   6            0.24859 g    -0.000004 se       0.0926 t       2.68
 nr   7            1.00515 g    -0.000003 se       0.1878 t       5.35
 nr   8            1.15976 g     0.000005 se       0.2115 t       5.48
 nr   9            1.14253 g    -0.000006 se       0.1967 t       5.81
 nr  10            9.89560 g    -0.000009 se       0.5076 t      19.50
 nr  11            9.70721 g     0.000001 se       0.5107 t      19.01
 nr  12           10.14058 g     0.000003 se       0.5867 t      17.28
 nr  13            9.94585 g     0.000002 se       0.5695 t      17.46
 nr  14            0.04708 g    -0.000420 se       0.0083 t       5.65
 nr  15            0.04881 g     0.000075 se       0.0084 t       5.82
 nr  16            0.04215 g     0.000165 se       0.0096 t       4.37
 nr  17            0.04530 g     0.000152 se       0.0093 t       4.85

 Logl        -324.192 -2xLogL         648.385 Inform(NPSOL)  0

title example  simulated

Model  1 of  1

S parameters
    1.000
    1.005
    1.160
    1.143

R parameters
    0.861
    0.000    0.884
    0.000    0.000    1.102
    0.000    0.000    0.000    0.831

H parameters
    0.769

Q parameters
    0.249


P parameters
  100.000

d parameters
    9.896    9.707   10.141    9.946

Z parameters
    0.000    0.047
    0.000    0.049
    0.000    0.042
    0.000    0.045

G parameters
    1.000

P(t|t) error cov
    0.122

start: 16 44 42 end: 16 44 48 date: 30  6 2003

start getse: 16 44 46 end getse: 16 44 48
```

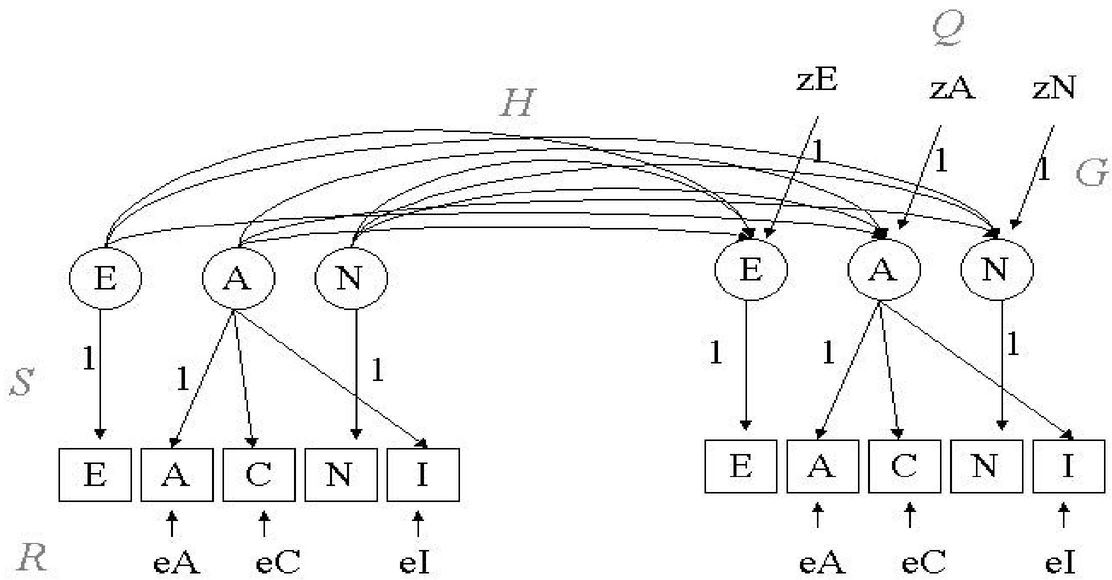
One worked example, data courtesy of Peter van Rijn.

We have three subjects, measured at 15 consecutive day. Variables are the big 5 dimensions (Extra,Agree,Consc,Neur,Intel).

Here are the data of the three cases:

```
                                                annotation
15                                              T=15
0 0 0                                           states at t=0
70.00 84.00 67.00 44.00 76.00                   the data T x N_of_variables
73.00 77.00 70.00 48.00 77.00
73.00 73.00 69.00 41.00 72.00
69.00 76.00 65.00 43.00 68.00
67.00 72.00 67.00 49.00 72.00
69.00 76.00 63.00 48.00 70.00
73.00 74.00 66.00 47.00 69.00
72.00 75.00 66.00 47.00 72.00
71.00 72.00 62.00 51.00 68.00
69.00 73.00 66.00 48.00 69.00
74.00 74.00 65.00 51.00 71.00
71.00 72.00 58.00 48.00 73.00
70.00 75.00 62.00 49.00 70.00
75.00 76.00 63.00 53.00 72.00
71.00 72.00 65.00 50.00 73.00

15
0 0 0
68.00 75.00 63.00 56.00 76.00
72.00 76.00 62.00 54.00 73.00
68.00 72.00 57.00 57.00 72.00
70.00 70.00 54.00 60.00 70.00
71.00 71.00 58.00 55.00 68.00
71.00 69.00 59.00 56.00 69.00
71.00 71.00 57.00 59.00 69.00
73.00 65.00 55.00 60.00 68.00
72.00 70.00 60.00 57.00 69.00
69.00 71.00 60.00 61.00 71.00
72.00 69.00 59.00 53.00 74.00
71.00 69.00 60.00 59.00 69.00
68.00 67.00 59.00 59.00 67.00
69.00 70.00 60.00 59.00 69.00
70.00 69.00 60.00 60.00 69.00

15
0 0 0
41.00 86.00 78.00 51.00 75.00
41.00 83.00 81.00 62.00 77.00
46.00 87.00 76.00 48.00 76.00
47.00 78.00 78.00 40.00 74.00
44.00 79.00 76.00 49.00 72.00
46.00 79.00 80.00 50.00 74.00
45.00 75.00 76.00 53.00 71.00
41.00 78.00 74.00 53.00 71.00
44.00 79.00 76.00 54.00 73.00
45.00 82.00 74.00 54.00 68.00
45.00 78.00 71.00 56.00 74.00
50.00 78.00 77.00 53.00 71.00
45.00 80.00 76.00 58.00 70.00
44.00 79.00 80.00 53.00 71.00
45.00 84.00 76.00 52.00 73.00
```

Due to mean differences, these three subject cannot be viewed as an emsemble. I therefore specify a nm=3 model, with within each of the nm model the means are estimated independently. We fit the following model:



The grey capital letters correspond to some of the model matrices:
Q is the diagonal covariance matrix of the innovation terms z;
G is the identitiy matrix
R is the covariance matrix of the residual terms e in the measurement model
S is the matrix this factor loadings of the observed variables on the latent variables (but 1-to-1 relationships for E and N)
H is the matrix with transition paramters.

Here is the input file:

```
title

nm=3 se=yes it=1000
mo=1 ny=5 ne=3 nx=0
df=subj1 rf=no ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=0 P=1 G=1 Z=0


R fi
0
0 0
0 0 0
0 0 0 0
0 0 0 0 0
R fr
0
0 1
0 0 2
0 0 0 0
0 0 0 0 3

H fi
0 0 0
0 0 0
0 0 0
```

```
H fr
10 11 12
13 14 15
16 17 18

Q fi di
0 0 0
Q fr
21 22 23

P fi
.5
0 .5
0 0 .5
P fr
0
0 0
0 0 0


G fi
1 0 0
0 1 0
0 0 1
G fr
0 0 0
0 0 0
0 0 0

d fi
0 0 0 0 0
d fr
41 42 43 44 45

S fi
1  0 0
0  1 0
0  0 0
0  0 1
0  0 0
S fr
0  0 0
0  0 0
0  32 0
0  0 0
0  33 0

mo=2 ny=5 ne=3 nx=0
df=subj2 rf=no ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=0 P=1 G=1 Z=0

G fi
1 0 0
0 1 0
0 0 1
G fr
0 0 0
0 0 0
0 0 0

R fi
0
0 0
0 0 0
0 0 0 0
0 0 0 0 0
R fr
0
0 1
0 0 2
0 0 0 0
0 0 0 0 3

H fi
0 0 0
0 0 0
0 0 0
```

```
H fr
10 11 12
13 14 15
16 17 18

Q fi di
0 0 0
Q fr di
21 22 23

P fi
.5
0 .5
0 0 .5
P fr
0
0 0
0 0 0


d fi
0 0 0 0 0
d fr
51 52 53 54 55

S fi

1  0 0
0  1 0
0  0 0
0  0 1
0  0 0
S fr
0  0 0
0  0 0
0  32 0
0  0 0
0  33 0

mo=3 ny=5 ne=3 nx=0
df=subj3 rf=no ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=0 P=1 G=1 Z=0

G fi
1 0 0
0 1 0
0 0 1
G fr
0 0 0
0 0 0
0 0 0

R fi
0

0 0
0 0 0
0 0 0 0
0 0 0 0 0
R fr
0
0 1
0 0 2
0 0 0 0
0 0 0 0 3

H fi
0 0 0
0 0 0
0 0 0

H fr
10 11 12
13 14 15
16 17 18
```

```
Q fi di
0 0 0
Q fr di
21 22 23

P fi
.5
0 .5
0 0 .5
P fr
0
0 0
0 0 0


d fi
0 0 0 0 0
d fr
61 62 63 64 65

S fi
1  0 0
0  1 0
0  0 0
0  0 1
0  0 0
S fr
0  0 0
0  0 0
0  32 0
0  0 0
0  33 0

st
 4 4 4
 .5 .1 .1
 .1 .5 .1
 .1 .1 .5
 5 1 5
 1 1
        44.60    80.33    76.60    52.40    72.67
        70.33    70.27    58.87    57.67    70.20
        71.13    74.73    64.93    47.80    71.47
lb
 0 0 0
-1 -1 -1
 -1 -1 -1
-1 -1 -1
 .0 .001 .0
  -10 -10
 0 0 0 0 0
 0 0 0 0 0
 0 0 0 0 0
ub
 100 100 100
1 1 1
1 1 1
1 1 1
 100 100 100
 100 100 100
 100 100 100
  100 100
 100 100 100 100 100
 100 100 100 100 100
 100 100 100 100 100
```

Note that I have chosen P to be very small. This is because large values of P, while more natural, result in computational problems due to the very short series (T=15).

Here is the output

```
max nm= 3 nt=5000 ns=  10 ny=30 nx= 5 ne=30 npar=400
```

```
Read from input file

title
nm=3 se=yes it=1000

mo=1 ny=5 ne=3 nx=0
df=subj1 rf=no ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=0 P=1 G=1 Z=0

mo=2 ny=5 ne=3 nx=0
df=subj2 rf=no ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=0 P=1 G=1 Z=0

mo=3 ny=5 ne=3 nx=0
df=subj3 rf=no ns=1 mi=-999
S=1 R=1 H=1 Q=1 d=1 c=0 P=1 G=1 Z=0

                          ===================
                          MKFv1     Nov. 2002
                          ===================

title

Model  1 of  3

S fr parameters (nonzero)
   0   0   0
   0   0   0
   0  16   0
   0   0   0
   0  17   0

R fr parameters (nonzero)
   0
   0   1
   0   0   2
   0   0   0   0
   0   0   0   0   3

H fr parameters (nonzero)
   4   5   6
   7   8   9
  10  11  12

Q fr parameters (nonzero) - diagonal
  13  14  15

d fr parameters (nonzero)
  18  19  20  21  22

Model  2 of  3

S fr parameters (nonzero)
   0   0   0
   0   0   0
   0  16   0
   0   0   0
   0  17   0

R fr parameters (nonzero)
   0
   0   1
   0   0   2
   0   0   0   0
   0   0   0   0   3

H fr parameters (nonzero)
   4   5   6
   7   8   9
  10  11  12

Q fr parameters (nonzero) - diagonal
  13  14  15

d fr parameters (nonzero)
  23  24  25  26  27
```

```
Model  3 of  3

S fr parameters (nonzero)
   0   0   0
   0   0   0
   0  16   0
   0   0   0
   0  17   0

R fr parameters (nonzero)
   0
   0   1
   0   0   2
   0   0   0   0
   0   0   0   0   3

H fr parameters (nonzero)
   4   5   6
   7   8   9
  10  11  12

Q fr parameters (nonzero) - diagonal
  13  14  15

d fr parameters (nonzero)
  28  29  30  31  32

DATA SUMMARY

MODEL  1 of  3 NY= 5 NX= 0 NE= 3 Ncases=   1 START=   1 END=   1

CASE    1    T=  15 N of T missing=   0 datafile subj1

   State_0    0.00   0.00   0.00
   var          1        2        3        4        5
   %miss   0.0000   0.0000   0.0000   0.0000   0.0000
   mean     44.60    80.33    76.60    52.40    72.67
   var       5.31    10.49     6.24    22.37     5.42
   std       2.30     3.24     2.50     4.73     2.33
   min      41.00    75.00    71.00    40.00    68.00
   max      50.00    87.00    81.00    62.00    77.00

MODEL  2 of  3 NY= 5 NX= 0 NE= 3 Ncases=   1 START=   2 END=   2

CASE    2    T=  15 N of T missing=   0 datafile subj2

   State_0    0.00   0.00   0.00
   var          1        2        3        4        5
   %miss   0.0000   0.0000   0.0000   0.0000   0.0000
   mean     70.33    70.27    58.87    57.67    70.20
   var       2.49     7.00     5.32     5.56     5.89
   std       1.58     2.64     2.31     2.36     2.43
   min      68.00    65.00    54.00    53.00    67.00
   max      73.00    76.00    63.00    61.00    76.00

MODEL  3 of  3 NY= 5 NX= 0 NE= 3 Ncases=   1 START=   3 END=   3

CASE    3    T=  15 N of T missing=   0 datafile subj3

   State_0    0.00   0.00   0.00
   var          1        2        3        4        5
   %miss   0.0000   0.0000   0.0000   0.0000   0.0000
   mean     71.13    74.73    64.93    47.80    71.47
   var       4.52     8.86     8.46     9.36     6.52
   std       2.12     2.98     2.91     3.06     2.55
   min      67.00    72.00    58.00    41.00    68.00
   max      75.00    84.00    70.00    53.00    77.00


 ML parameter estimates
 nr   1            3.73787 g    -0.000001 se       0.9308 t       4.02
 nr   2            5.28120 g     0.000002 se       1.1922 t       4.43
 nr   3            2.68485 g     0.000000 se       0.7395 t       3.63
 nr   4            0.54846 g     0.000044 se       0.2230 t       2.46
 nr   5           -0.18268 g     0.000011 se       0.1555 t      -1.18
 nr   6            0.21345 g     0.000014 se       0.1233 t       1.73
```

nr	7	-0.44093	g	0.000008	se	0.1216	t	-3.63
nr	8	0.76250	g	-0.000020	se	0.1054	t	7.23
nr	9	0.11590	g	-0.000002	se	0.1094	t	1.06
nr	10	-0.16787	g	-0.000020	se	0.2435	t	-0.69
nr	11	-0.70487	g	0.000054	se	0.2654	t	-2.66
nr	12	0.02314	g	0.000009	se	0.1527	t	0.15
nr	13	5.38001	g	0.000002	se	1.5174	t	3.55
nr	14	0.66878	g	-0.000004	se	0.6911	t	0.97
nr	15	9.43727	g	-0.000011	se	2.0828	t	4.53
nr	16	0.52140	g	-0.000055	se	0.1870	t	2.79
nr	17	0.79796	g	-0.000027	se	0.1639	t	4.87
nr	18	41.93236	g	-0.000003	se	1.5464	t	27.12
nr	19	83.56521	g	0.000002	se	1.5271	t	54.72
nr	20	78.28511	g	0.000010	se	1.0555	t	74.17
nr	21	50.54019	g	0.000014	se	1.4383	t	35.14
nr	22	75.24557	g	-0.000012	se	1.2055	t	62.42
nr	23	66.56459	g	0.000007	se	1.6692	t	39.88
nr	24	74.96289	g	0.000004	se	1.5401	t	48.67
nr	25	61.31529	g	0.000004	se	1.2040	t	50.93
nr	26	55.10621	g	-0.000013	se	1.8148	t	30.36
nr	27	73.94739	g	-0.000009	se	1.2715	t	58.16
nr	28	66.66162	g	0.000008	se	2.0062	t	33.23
nr	29	80.23622	g	0.000003	se	1.6611	t	48.30
nr	30	67.80255	g	0.000005	se	1.2701	t	53.38
nr	31	44.91112	g	-0.000006	se	1.9880	t	22.59
nr	32	75.85773	g	-0.000007	se	1.2745	t	59.52

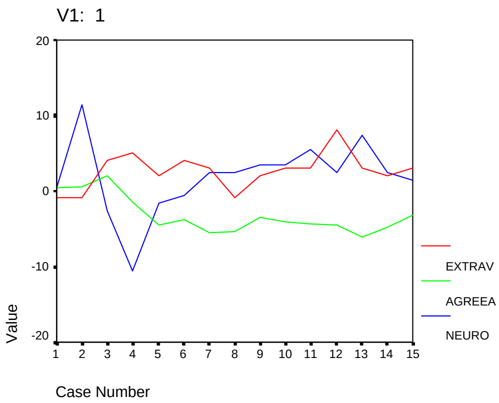
```
 Logl        -342.374 -2xLogL         684.748 Inform(NPSOL)  0

title

Model  1 of  3

S parameters
    1.000    0.000    0.000
    0.000    1.000    0.000
    0.000    0.521    0.000
    0.000    0.000    1.000
    0.000    0.798    0.000

R parameters
    0.000
    0.000    3.738
    0.000    0.000    5.281
    0.000    0.000    0.000    0.000
    0.000    0.000    0.000    0.000    2.685

H parameters
    0.548   -0.183    0.213
   -0.441    0.763    0.116
   -0.168   -0.705    0.023

Q parameters - diagonal
    5.380    0.669    9.437

P parameters
    0.500
    0.000    0.500
    0.000    0.000    0.500

d parameters
   41.932   83.565   78.285   50.540   75.246

G parameters
    1.000    0.000    0.000
    0.000    1.000    0.000
    0.000    0.000    1.000

P(t|t) error cov
    0.000
    0.000    0.658
    0.000    0.000    0.000


Model  2 of  3

S parameters
```

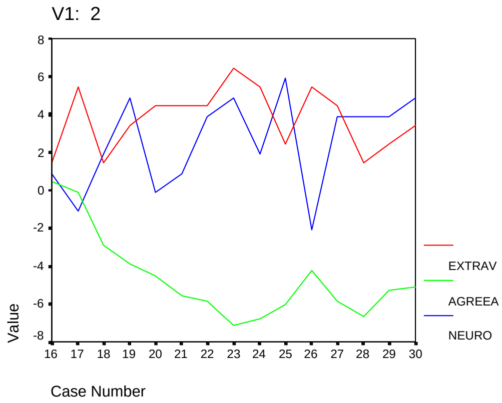
```
    1.000    0.000    0.000
    0.000    1.000    0.000
    0.000    0.521    0.000
    0.000    0.000    1.000
    0.000    0.798    0.000

R parameters
    0.000
    0.000    3.738
    0.000    0.000    5.281
    0.000    0.000    0.000    0.000
    0.000    0.000    0.000    0.000    2.685

H parameters
    0.548   -0.183    0.213
   -0.441    0.763    0.116
   -0.168   -0.705    0.023

Q parameters - diagonal
    5.380    0.669    9.437

P parameters
    0.500
    0.000    0.500
    0.000    0.000    0.500

d parameters
   66.565   74.963   61.315   55.106   73.947

G parameters
    1.000    0.000    0.000
    0.000    1.000    0.000
    0.000    0.000    1.000


P(t|t) error cov
    0.000
    0.000    0.658
    0.000    0.000    0.000


Model  3 of  3

S parameters
    1.000    0.000    0.000
    0.000    1.000    0.000
    0.000    0.521    0.000
    0.000    0.000    1.000
    0.000    0.798    0.000

R parameters
    0.000
    0.000    3.738
    0.000    0.000    5.281
    0.000    0.000    0.000    0.000
    0.000    0.000    0.000    0.000    2.685

H parameters
    0.548   -0.183    0.213
   -0.441    0.763    0.116
   -0.168   -0.705    0.023

Q parameters - diagonal
    5.380    0.669    9.437

P parameters
    0.500
    0.000    0.500
    0.000    0.000    0.500

d parameters
   66.662   80.236   67.803   44.911   75.858

G parameters
    1.000    0.000    0.000
    0.000    1.000    0.000
    0.000    0.000    1.000
```

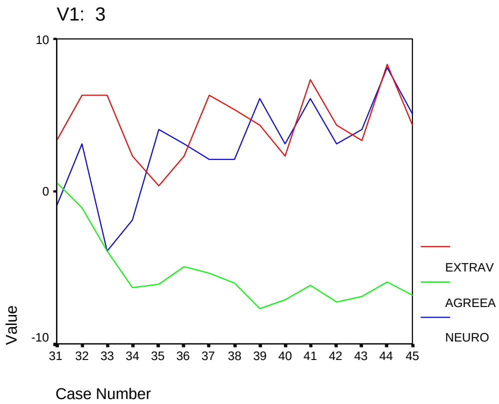
```
P(t|t) error cov
    0.000
    0.000    0.658
    0.000    0.000    0.000


start: 10 38 26 end: 10 38 41 date:   4  2 2005

start getse: 10 38 35 end getse: 10 38 41
```

Although we estimate 32 parameter, 15 of these correspond to the observed means, and thus are easy to estimate. The remaining 17 are contrained to be equal over the 3 cases. Clearly not all parameters are significant, so one would prune the model (at the risk of chance capitilization).

Here are the estimates states for three cases (change *rf=no* to *rf=filename* to obtain these):







Subroutines & glossary [not important].

structure	function	calls
MKF_ML*	main	readi errcheck readd nsand writer fun badrep npoptn npsol getse writep writed setz setnp sumdat
setz	sets arrays to zero	
readi*	reads input script	getvn
- getvn	finds keywords in a character string	-
errcheck	reports errors	geterror
- geterror	gets error messages	-
readd	reads input data from file XXXX (file df=XXXX)	-
nsand	determines begin and end of groups in data file	-
writer	writes model specification as read from input file	-
fun*°	loglikelihood function	kf
kf*° (kfmis, kfmis2)	kalman filter	mxv1 mxm1 mxm2 cholinv
mvx1, mxm1,mxm2	matrix, vector multiplication	-
cholinv	invert pos def symmetric matrix	cholesk
cholesk	cholesky decomp. of pos def symmetric matrix	-
badrep	reports non-fatal errors	-
npoptn	set npsol options	(see npsol doc)
npsol	optimizer	many (see npsol doc) objfun
objfun	likelihood function and gradients	fun (see above) grd
fun*°	calculates loglikelihood function	kf (kfmis, kfmis2)
grd*°	calculates gradients of loglikelihood function	kf (see above) dfk (dkfmis, dkfmis2)
dkf*° (dkfmis,dkfmis2)	calculates gradients	mxv1 (see above) mxm1 mxm2
getse	calculates standard errors	fdh cholinv (see above)
fdh	finite difference approximation of hessian	fun (see above)
writep	writes parameter estimates, st errs, etc.	-
writed	writes state vector to XXXXX (rf=XXXXX)	-
setnp	set some npsol parameters	-
sumdat	calculates summary statistics	-

*these routines include parameter statements:

MAXNY	max length of vector of observations
MAXNX	max length of vector of exogenous (fixed) regressors
MAXNE	max length of vector of states
MAXD	max(maxny, maxne)
MAXNM	number of model (groups)
MAXNT	max length of timeseries
MAXNS	max number of subjects, or cases.
NPARM	number of parameter matrices / vectors (= 9)
MAXP	max number of parameters

°these routines contain the actual implementation of Harvey p. 144ff.
kfmis (dkfmis) and kfmis2 (dkfmis2) include provision for missing data

Arrays & parameters

Parameters

MAXNY	max length of vector of observations
MAXNX	max length of vector of fixed regressors
MAXNE	max length of vector of states
MAXD	max(maxny, maxne)
MAXNM	number of model (groups)
MAXNT	max length of timeseries
MAXNS	max number of subjects, or cases.
NPARM	number of parameter matrices / vectors (= 9)
MAXP	max number of parameters

These {or a subset) occur at six places in the source code. See subroutines MKF_ML (main), READI, GRD, FUN, DKFMIS, DKFMIS2, DKF, KFMIS, KFMIS2, KF, SUMDAT.

Arrays

TS(maxns,maxnt,maxny)	observed timeseries	
EXO(maxns,maxnt,maxny)	fixed regressors	
xest(maxns,maxnt,maxne)	estimates states	
xst(maxns,maxne)	**a**[0]	
xmat(maxnm,nparm,maxd,maxd)	contains parameter matrices/vectors	
xmat(maxnm,1,maxny,maxne)	**S**	full
xmat(maxnm,2,maxny,maxny)	**R**	symmetric
xmat(maxnm,3,maxne,maxne)	**H**	full
xmat(maxnm,4,maxne,maxne)	**Q**	symmetric
xmat(maxnm,5,maxne,maxne)	**P**	symmetric cov matrix of **a**[0]
xmat(maxnm,6,maxny,1)	**d**	vector
xmat(maxnm,7,maxne,1)	**c**	vector
xmat(maxnm,8,maxny,maxnx)	**Z**	full
xmat(maxnm,9,maxny,maxnx)	**G**	full
idim(maxnm,nparm,2)	dimensions of **S, R, H, Q, P, d**, **c, Z, G**	
pend(maxnm,maxne,maxne)	estimated error cov matrix of **a**[t] at end of optimization	
imat(maxnm,nparm,maxd,maxd)	indicator matrices, contain zeros or integer each integer is associated with a free parameter	
nt(maxns)	length of timeseries of each case	
ns(maxnm)	number of subjects in each model (or group)	
nns(maxnm,2)	start and end of each case in group	