FREM technical description

2013-06-26

Introduction

The frem script creates a set of models to gradually include random effects on "everything" to simultaneously account for all parameter-covariate relations. [TODO Description of frem methodology should be added.] If option -vpc is set then script will prepare model that can be run with vpc.

Example
frem -covariates run1.mod

For BOV the script defines a variable BOV<par> for each parameter, and then BOV<par> is set equal to different ETAs depending on the occasion. Same for BOV<covariate>. See details below. This method is safer to implement than having switch variables for occasions that are 1 or 0 depending on the occasion, and then use a long sum of products between switch variables and ETAs.

Input and options

frem-specific input

A model file is required on the command-line. Unless option -model3 is set (see below) the input model code must not include covariates that are listed in options -invariant or -time_varying. Also, the input model must not include BOV on any parameters, only BSV is allowed. For each parameter listed with option -parameters the input model must have a line

<par>=TV<par> (BSV-expression)

If there is no BSV for the parameter, i.e. the value is equal to the typical value, the user must still include an expression with 0 in the place of the BSV eta, using the desired form (additive, proportional, exponential,..). This is necessary to enable PsN to automatically add the BOV code. The 0 must be enclosed in parentheses, otherwise the frem program will not recognize it. Example with placeholder for additive inclusion of BOV:

<par>=TV<par> + (0)

If option -model3 is set then it is assumed that the input model file is a Model 3 created earlier by the frem script, and then other assumptions about the contents of the model are made. See below for a detailed description of Model 3.

frem options
-invariant=<list>	A comma-separated list of covariates that are time invariant. Optional if -time_varying is given, not allowed if -model3 is set, otherwise required.

	Names used in $INPUT.
-time_varying=<list>	A comma-separated list of time-varying covariates. Optional if -invariant is given, not allowed if -model3 is set, otherwise required. Names used in $INPUT.
-occasion=name	Default OCC. Needed if -time_varying is set or both -model3 and -bov_parameters, otherwise ignored. The name of the column that defines occasions. Name used in $INPUT, not data set.
-parameters=<list>	Required if -time-varying is set, otherwise not allowed. A comma-separated list of model parameters that should have between-occasion variability added.
-dv=name	Default is DV. The name of the dependent variable. Name used in $INPUT.
-vpc	Default not set. If set then script will prepare frem model that can be run with the vpc script (separate PsN call). If -model3 is set then this option will be set automatically.
-model3	Default not set. If set then it is assumed that the input model is of type Model 3 created earlier by this script. Then frem will only create vpc model. In this case some of options above are not allowed, but need bsv_etas and some others instead, see below.
-check	Set by default, disable with -no-check. Run safety check after data set 2 generation.
-estimate=number	Optional, default not defined. The number (0, 1, 2 or 3) of the last model to estimate in the frem sequence. All models with a higher number will still be created, but not estimated. If option model3 is not set and -estimation >= 0, estimation of Model 0 (the input model) will be skipped if a lst-file for the input model is found, i.e. a file with same stem but extension .lst instead of .mod. If option model3 is set and -estimation>=3, estimation of Model 3 (the input model when option model3 is set) will be skipped if a lst-file is found, i.e. a file with same stem but extension .lst instead of .mod. TODO Functionality to update initials of individual models with better estimates if some models are run at a later stage.
-bsv_parameters=N	Required if -model3 is set, otherwise not allowed. The number of model parameters (not covariates) that have BSV etas. (This number should be equal to the number of etas in Model 0/input model in the frem run that generated the input model3)
-N_invariant=N	Required if -model3 and -bov_parameters are both set, otherwise not allowed. The number of invariant covariates that have BSV etas. (This number should be equal to the number of invariant covariates given with option -invariant to the frem run that generated the input model3)
-bov_parameters=N	Optional if -model3 is set, otherwise not allowed. The number of model parameters that have BOV etas. (This number should be equal to the number of parameters given with option -parameters to the frem run that generated the input model3)
-type=COL	Default FREMTYPE. Only allowed if -model3 is set. The name of the column that distinguishes between original and covariate observations. The

value in the FREMTYPE column must be 0 for original observations and larger than 0 for covariate observations.

Some general PsN-options which are interesting in combination with frem

For a complete list of common options see common_options_defaults_versions.pdf, or psn_options -h on the command-line.

-directory=name   Name of run directory

-c   A
-o   R.

Model 0, Dataset 0

The input model, input dataset.

Dataset 2

Create a new data set.
1) Determine a name for the FREMTYPE column to be added to the input data set. TODO: If $INPUT, $PK, $PRED or $ERROR already contains a variable called FREMTYPE then check if FREMTYPE1 is available, and so on. In following steps assume name is FREMTYPE.
2) Run a dummy NONMEM model where FREMTYPE=0 to filter the original data on any IGNORE/ACCEPT and to get a new data file from $TABLE where the column headers are the same as in $INPUT but a FREMTYPE column is added with all zeros as values. $TABLE NOPRINT ONEHEADER NOAPPEND. Add MDV if not EVID or MDV in data set. Do not need to use RFORMAT in $TABLE since problem with .EQ. and string matching only occurs with IGNORE/ACCEPT.
3) In filtered dataset: Find EVID and/or MDV columns. If neither found then stop with error message. Find DV column. Find FREMTYPE column. Find OCC column, if there are time-varying covariates listed with the time-varying option. For each invariant and time-varying covariate determine a unique FREMTYPE value > 0. FREMTYPE=0 is reserved for original observations.

Then loop over each individual:
First: Copy first line for individual. Loop over each invariant covariate: Add new line as first data set line for individual. Set DV value to covariate value on copied line. Set FREMTYPE to type-value for this covariate. If MDV present then MDV=0 if have covariate value, MDV=1 if value is missing. If EVID present then EVID=0 if non-missing covariate value, otherwise EVID=2. Store non-missing covariate values in array, one array per covariate, to be able to compute medians and variance-covariance matrix later to be used as initial estimates in $THETA and $OMEGA.
Second: Loop over each occasion (if OCC defined). Copy first line for this occasion. Warn if multiple covariate values within the data records of one occasion. Inner loop over each time-varying covariate: Add new line as first data set line for this occasion. Set DV value to covariate value for this occasion.

Set FREMTYPE to type-value for this covariate. If MDV present then MDV=0 if have covariate value, otherwise MDV=1. If EVID present then EVID=0 if have covariate value, otherwise EVID=2. Store non-missing time-varying covariate values to compute median over occasions for this individual, median is one scalar per covariate. Then store non-missing medians in array, one array per time-varying covariate, to compute median of medians later for $THETA, and variance-covariance matrix to be used in $OMEGA.

<u>end loop over individuals</u>
   4) Compute medians of invariant and time-varying covariates. Also compute variance-covariance matrix of invariant covariates, and of medians (over occasions) for time-varying. Two separate variance-covariance matrices. If missing values then just skip.

This dataset is used for Model 2 and all consecutive models.

Data check model, Dataset 2

Copy Model 0.

DATA changes
Set $DATA to dataset 2. Skip old IGNORE/ACCEPT. Set IGNORE=@ IGNORE=(FREMTYPE.GT.0)

INPUT changes
Append extra items to $INPUT: MDV(possibly) and FREMTYPE.

For safety check, done if option -check is set: Estimate Data check model. OFV should be identical to Model 0 OFV.

Model 1

This filling in of $OMEGA is only done if there are invariant covariates listed with the -invariate option.
Copy the input model and change $OMEGA from whatever form in Model 0 to a single large BLOCK. Initials that are non-zero in Model 0 are copied from there or taken from Model 0 lst-file, if present. Off-diagonal elements that are new to Model 1 are set to small non-zero values that preserve strict diagonal dominance OR compute ETA-covariances from estimation output phi file if available.
The large $OMEGA block of Model 1 is called the BSV_par block.

Model 2-all covariates

SETUP
Count number n_theta of already present THETAS in Model 0. Count number bsv_parameters of already present ETAs. (equal to dimensionality of total $OMEGAs in Model 0). Count number of n_eps already present (dimensionality of $SIGMA in Model 0), set epsnum=n_eps+1.
Build Model 2 based on Model 1, which is only different from Model 0 if option -invariant is set (a

filled BSV_par block).

DATA changes
Set $DATA to dataset 2. Skip old IGNORE/ACCEPT. Set IGNORE=@.

INPUT changes
Append extra items to $INPUT: MDV(possibly) and FREMTYPE.

OMEGA changes
BSV_par: Already have full BSV_par block from Model 1.
BSV_cov: Add a full BLOCK(N_invariant) where initials is variance-covariance matrix of invariant covariates computed during creation of dataset.
At each line, possibly add comment with covariate name to be used as label for ETA in PsN
Then, provided have time-varying covariates:
BOV_par_occ1: Add one full BLOCK(bov_parameters) where initials is: diagonal 0.01 and off-diagonals small enough to make strictly diagonally dominant
At each line, possibly add comment with BOV<parameter> to be used as label for ETA in PsN
BOV_par_occ2-end: Then, for each occasion that is not the first occasion add BLOCK SAME.
BOV_cov_occ1: Add one full BLOCK(N_time-var) where initials is variance-covariance matrix of time-varying covariates computed during creation of dataset
At each line, possibly add comment with covariate name to be used as label for ETA in PsN
BOV_cov_occ2-end: Then, for each occasion that is not the first occasion add BLOCK SAME.

SIGMA changes
In $SIGMA add
$SIGMA 0.0000001 FIX; EPSCOV

THETA changes
Loop over all covariates (invariant plus time-varying), j=1...N_covtot, add lines
$THETA init ; TVcovname
where init is median of covariate computed during creation of new data set. Store mapping Tvcovname – THETA(n_theta+j). Possibly set boundaries to max and min for covariate, in that case need to find that during generation of dataset 2.

PK or PRED changes at beginning, A
In $PK or $PRED, whichever is present:
At the very beginning (unless have FREM-ANCHOR analogue to scm- anchor) the following code:
Loop over invariant covariates j=1...N_invariant, add lines
BSV<cov> = ETA(bsv_parameters+j)
Loop over parameters, initialize BOV PK (provided there are any time-varying cov)
BOV<parameter>=0
Loop over occasions k=1...N_occ , one IF block per occasion
IF (OCC .EQ. k) THEN
[inner loop over parameters j=1...bov_parameters]

BOV<parameter> = ETA(bsv_parameters+N_invariant+(k-1)*bov_parameters+j)
END IF
Loop over time-varying covariates, initialize BOV tcov
BOV<covariate>=0
Loop over occasions k=1...N_occ , one IF block per occasion
IF (OCC .EQ. k) THEN
[inner loop over time-varying covariates j=1...N_time-var]
BOV<covariate> = ETA(bsv_parameters+N_invariant+N_occ*Nparams+(k-1)*N_time-var+j)
END IF

PK or PRED changes at beginning, B

In $PK or $PRED (continuing):
Loop over parameters (provided there are any time-varying cov)
Locate line with
<par> = TV<par> (expression defining form of BSV/IIV relation)
Raise error if not found, or if multiple instances are found. Use old lin-scm code to figure out which
form (additive, exp, prop) of BSV relation. Require user to define placeholder with right form if
<par>=TV<par>, i.e, no BSV eta at all. Add BOV<parameter> at appropriate place in
<par>=TV<par>... expression.

PRED or ERROR changes at end

In $PRED or $ERROR, whichever is present:
At the very end
Loop over invariant covariates, j=1...N_invariant, use THETA-covname mapping, add lines
Y<j> = THETA(cov) + BSV<cov>
Loop over time-varying covariates, j=(N_invariant+1)...(N_invariant+N_time-var), use THETA-
covname mapping, add lines
Y<j> = THETA(cov) + BOV<cov>

Loop over all covariates j=1...(N_invariant+N_time-var).
For each cov find FREMTYPE integer k from mapping, add IF block code
IF (FREMTYPE .EQ.k) THEN
Y=Y<j>+EPS(epsnum)
IPRED=Y<j>
END IF

(do not define a case for FREMTYPE=0, original obs. Leave original code as it is for that)

Model 2-only time-varying

Use dataset but in $DATA set IGNORE for all FREMTYPE for invariant covariates. Build model based
on Model 0 (no filling in of BSV block). Add code and parameters according to Model 2 instructions
above but set the number of invariant covariates N_invariant to 0.

Model 3

This models have same number of ETAs as Model 2, but the BSV and BOV blocks are filled in, 1+N_occasions huge $OMEGA blocks and reordered ETAs. Take care of initials.

SETUP
Same as for Model 2 (build on Model 0 or 1)

DATA changes
Same as Model 2.
INPUT changes
Same as Model 2.

OMEGA changes
BSV_all: One full BLOCK(bsv_parameters+N_invariant) where bsv_parameters is total dimensionality of $OMEGAs in Model 0 and N_inv is number of invariant covariates.
Initials: EITHER from eta estimates from Model 2 (from .phi-file), compute variance-covariance matrix. OR same initials as in Model 2 plus small values to fill in full BLOCK.
At lines for covariates add comment with covariate name to be used as labels in PsN.

Then, provided have time-varying covariates:
BOV_all_occ1: Add one full BLOCK(bov_parameters+N_time-var) where bov_parameters is number of parameters set with -parameters option.
At each line add comment with name to be used as labels in PsN.
Initials: EITHER from eta estimates from Model 2, compute variance-covariance matrix (.phi-file, must rearrange ETAs to new order in Model 3). OR same initials as in Model 2 plus small values to fill in full BLOCK.
BOV_all_occ2-end: Then, for each occasion that is not the first occasion add $OMEGA BLOCK SAME.

SIGMA changes
Same as Model 2.

THETA changes
Same as Model 2.

PK or PRED changes at beginning A
Very similar to Model 2, but different arrangement of loops so that ETA numbering is different.
In $PK or $PRED, whichever is present:
At the very beginning (unless have FREM-ANCHOR analogue to scm- anchor) the following code:
Loop over invariant covariates j=1...N_inv, add lines
BSV<cov> = ETA(bsv_parameters+j)
Loop over parameters, initialize BOV PK (provided there are any time-varying cov)
BOV<parameter>=0

Loop over time-varying covariates, initialize BOV tcov
BOV<covariate>=0
Loop over occasions k=1...N_occ , one IF block per occasion
IF (OCC .EQ. k) THEN
[First inner loop, over parameters j=1...bov_parameters]
BOV<parameter> = ETA(bsv_parameters+N_inv+(k-1)*(bov_parameters+N_variant)+j)
[Second inner loop (still inside IF BLOCK, but outside first inner loop), over time-varying covariates j=1...N_variant]
BOV<covariate> = ETA(bsv_parameters+N_inv+(k-1)*(bov_parameters+N_variant)
+bov_parameters+j)
END IF

PK or PRED changes at beginning B
Same as Model 2.

PRED or ERROR changes at end
Same as Model 2.

ANNOTATION
Add comment lines with FREM-flag and info about bsv_parameters, n_theta, n_eps, N_invariant, N_time-var, bov_parameters. Occasion column. FREMTYPE-cov mapping. DV, Model 3.

Creating FREM vpc model

Done if option -vpc is given to FREM script. This part of the recipe describes modifications needed to FREM model and data set before running vpc. Definitions of model types (Model 2 and Model 3) and names of $OMEGA blocks are as in description above.

If a frem model file of type Model 3 is given on the command-line, then the frem script will skip generation of Model 1-3 above and go directly to this part. Then that input model must have form of Model 3, i.e. $OMEGA has one BSV_all block, one BOV_all_occ1 block for the first occasion and one BLOCK SAME for each subsequent occasion, see above for definitions of OMEGA block types. Assume FREMTYPE etc as in FREM description above.

Step 1: Obtain typical parameter values conditional on covariate info

Create frem_vpc1.mod

1. Take input model (Model 3). Update inits from estimation, if available. TODO: Option to run Model 3 if lst not available. Afterwards FIX $THETA and $OMEGA. Unfix all $SIGMA.

2. Set IGNORE=(FREMTYPE.GT.0) where FREMTYPE can be replaced using option -type.

(The FREMTYPE that is not a time-varying or invariant covariate. Use .GT. to ensure numerical comparison instead of string comparison in NONMEM.)

3. Remove $COV.

4. In $PK or $PRED, whichever is present:

Find all THETAs on right hand side where left hand is *not* Y<number> or Y. Store mapping THETA number and left hand parameter name. Assume at most one THETA in right hand expression.

Locate ETA(1) to ETA(bsv_parameters) where bsv_parameters is the number of BSV etas for parameters. This number is given as input to script or read from annotation as bsv_parameters in Model 0, assume numbering starts on 1 and consecutive. Figure out the mapping between ETAnumber and THETAnumber. The script shall handle link directly on same line as in

CL=THETA(2)*EXP(ETA(1)),

or indirectly via THETA-parameter mapping as in

TVCL=THETA(2)

CL=TVCL*EXP(ETA(1)).

These are the two cases in Ron's vpc HO. Assume no more complicated connections than that. If not find THETA number then raise error.

5. Add $TABLE with all undropped $INPUT variables + list of names of parameters that have BSV eta + ONEHEADER NOAPPEND NOPRINT FILE=conditional_tv.tab

6. Store sequence of headers in $TAB to be used in $INPUT in model frem_vpc2.mod, but in this list prepend parameter names with TV to separate from other params.

7. Run model with NM.

Step 2: Create new dataset which is original dataset + TVs

Take original dataset, add one column per TV-parameter. Take values from conditional_tv.tab which has the same number of lines as original dataset. This is different from Ron's method which uses FIRSTONLY and R script frem_vpc_transfer_tv.R.

outfile is comb_cond_tvs.csv

Step 3: Calculate parameter OMEGA conditional on covariate OMEGA

Create frem_vpc2.mod

1. Take input model (Model 3). Update inits from estimation of input model, if available.

2. FIX $SIGMA

3. DATA is comb_cond_tvs.tab

4. Set IGNORE=(FREMTYPE.GT.0), same as in Step 1.

5. Set $INPUT to sequence of headers from step 1 (contents of table file conditional_tv.tab but prepend TV for params)

6. Use THETA-parameter and ETA-THETA mappings from Step 1 to replace THETA(x) with TV<par> for parameters that have BSV ETA on them.

7. Read final estimated BSV_all (omega) block from lst-file from Step 1 and write to file om_bsv_block.csv. Compute conditional omega block as in frem_vpc_cond_om.R and write to om_bsv_block_cond.csv, see Appendix "Compute conditional omega block" below. In $OMEGA replace BSV_all block (Model 3) with BSV_par block as conditional omega block from om_bsv_block_cond.csv, plus one $OMEGA 0 FIX for each invariant covariate. The zeros are instead of a BSV_cov block.

8. For BOV???? Read final estimated BOV_full_occ1 omega block from lst-file from Step 1 and write to file om_bov_block.csv. Compute conditional omega block as in frem_vpc_cond_om.R and write to om_bov_block_cond.csv, see section "Compute conditional omega block" below.

9. For BOV???? In input model (Model 3) locate and modify section where BOV<parameter> and BOV<covariate> are defined. Set ETA numbering to have first N_occasion sets of ETAs for BOV parameters, this will give the same numbering as in Model 2 in FREM recipe. Set all BOV<covariate> equal to zero, no ETAs for them at all (alternatively define them exactly as in Model 2 and set all 0 FIX in $OMEGA). In $OMEGA remove BOV_full blocks for all occasions and instead set BOV_par_occ1 as conditional block from om_bov_block_cond.csv. Add one BLOCK SAME for each occasion that is not the first, BOV_par_occ2-end as in Model 2. No BOV_cov blocks at all since those ETAs have been removed.

Step 4: Prepare to run vpc

Remove or rename frem_vpc2.lst if present, do not want initial estimates to be changed in frem_vpc2.mod.

Now a regular vpc can be run on frem_vpc2.mod, but let user start that separately. Otherwise unhappy mix of options for vpc and frem scripts.

Appendix: Compute conditional omega block

frem_vpc_cond_om.R

Input is omega block from lst-file (or csv file om_block.csv)

(Detailed description TODO)

Ok to run on block from BOV omega?