Growth Regressions, Principal Components Augmented Regressions and Frequentist Model Averaging

Martin Wagner and Jaroslava Hlouskova

Forthcoming in: Journal of Economics and Statistics

Program runs from **main1.m**. Input data are from **Data.xls** and some comments are in sheet **Groups**. Variables *path_in* (the path where input data Data.xls are stored) and *path_out* (the path where results will be stored) need to be adjusted. For running the code, the kernel density estimation toolbox will be necessary which can be downloaded from

http://intarch.ac.uk/journal/issue1/beardah/kdeia8.html

The main results are stored in the structures. The ordering of struct arrays corresponds to the ordering of groups as in the text object named *names_groups* which appears in the final mat file (or starting from the 3rd line in the data file Data.xls); i.e., if we have a structure named *s* with arrays *s(1)* and *s(2)* and *names_groups*={Macro, Educations}, then information in *s(1)* corresponds to group Macro and in *s(2)* to Education.

Below is the description of output structures and their fields:

ESTIM

coefFMAaic: $1 \times (k_{11}+k_2+k_{12})$ vector of estimated model average coefficients with S-AIC weighting scheme

coefFMAbic:$1 \times (k_{11}+k_2+k_{12})$ vector of estimated model average coefficients with S-BIC weighting scheme

coefFMAequal: $1 \times (k_{11}+k_2+k_{12})$ vector of estimated model average coefficients with equal weighting scheme

coefFMAmma: $1 \times (k_{11}+k_2+k_{12})$ vector of estimated model average coefficients with MMA weighting scheme

coefOLSall: $2^{k12} \times (k_{11}+k_2+k_{12})$ matrix of estimated coefficients from all OLS regressions (coefficients of estimated factors are listed as well)

coefOLS: $2^{k12} \times (k_{11}+k_{12})$ matrix of estimated coefficients from all OLS regressions (coefficients of estimated factors are not listed)

mean: $1 \times (k_{11}+k_{12})$ vector of means of coefficients corresponding to variables in X_{11} and X_{12}

kurt: $1 \times (k_{11}+k_{12})$ vector of kurtosis of coefficients corresponding to variables in X_{11} and X_{12}

skew: $1 \times (k_{11}+k_{12})$ vector of skewness of coefficients corresponding to variables in X_{11} and X_{12}

std: $1 \times (k_{11}+k_{12})$ vector of standard deviations of coefficients corresponding to variables in X_{11} and X_{12}

INFERENCE

bounds_low: $(k_{11}+k_2+k_{12})*(3+flag\_MMA) \times nsiglevel$ matrix of lower bounds of confidence interval (based on the 2-stage confidence simulation procedure) for conservative coverage above 100*(1-sig_level_FMA) sign. level. First $k_{11}+k_2+k_{12}$ rows correspond to lower bounds for model average coefficients with equal weighting scheme, second $k_{11}+k_2+k_{12}$ rows correspond to lower bounds for model average coefficients with S-AIC weighting scheme, third $k_{11}+k_2+k_{12}$ rows correspond to bounds for model average coefficients with B-AIC weighting scheme and forth $k_{11}+k_2+k_{12}$ rows correspond to lower bounds for model average coefficients with MMA weighting scheme. *nsiglevel* indicates the number of significance levels based on which is the inference performed. E.g., if we deal with 2 significance levels (*nsiglevel*=2) such that first is performed on 5% sig. level and the second on 10% sig. level, then first column corresponds to the lower bound based on 5% sig. level and the second column corresponds to the lower bound based on 10% sig. level. Sig. levels for confidence intervals in FMA inference is given in *sig_level_FMA*.

bounds_up: $(k_{11}+k_2+k_{12})*(3+flag\_MMA) \times nsiglevel$ matrix of upper bounds of confidence interval (based on the 2-stage confidence simulation procedure) for conservative coverage above 100*(1-sig_level_FMA) sign. level. First $k_{11}+k_2+k_{12}$ rows correspond to upper bounds for model average coefficients with equal weighting scheme, second $k_{11}+k_2+k_{12}$ rows correspond to upper bounds for model average coefficients with S-AIC weighting scheme, third $k_{11}+k_2+k_{12}$ rows correspond to bounds for model average coefficients with B-AIC weighting scheme and forth $k_{11}+k_2+k_{12}$ rows correspond to upper bounds for model average coefficients with MMA weighting scheme. *nsiglevel* indicates the number of significance levels based on which is the inference performed. E.g., if we deal with 2 significance levels (*nsiglevel*=2) such that first is performed on 5% sig. level and the second on 10% sig. level, then first column corresponds to the upper bound based on 5% sig. level and the second column corresponds to the upper bound based on 10% sig. level. Sig. levels for confidence intervals in FMA inference is given in *sig_level_FMA*.

pvalFMA: $(k_{11}+k_2+k_{12}) \times (3+flag\_MMA)$ matrix of p-values for estimated model average coefficients with equal, S-AIC, B-AIC weighting scheme and MMA weighting scheme (if *flag_MMA*=0).

pvalOLS: $2^{k12} \times (k_{11}+k_2+k_{12})$ matrix of p-values of all coefficients and all OLS models.

OPTIM

exitflag: describes the exit condition of quadprog function (when calculating MMA weights); 1 - quadprog converged with a solution, 3 - change in objective function value smaller than the specified tolerance, 4 - local minimizer found, 0 - maximum number of iterations exceeded, -2 - no feasible point found, -3 - problem is unbounded, -4 - current search direction is not a descent direction; no further progress can be made, -7 Magnitude of search direction became too small; no further progress can be made. The problem is ill-posed or badly conditioned.

PCA

eigenval: k_{21} *x 1* vector of eigenvalues (for quantitative variables if *ngroups_PCA*=2)

eigenvald: k_{22} *x 1* vector of eigenvalues for dummy variables (if *ngroups_PCA*=2)

eigenvec: k_{21} *x* k_{21} matrix of eigenvectors (for quantitative variables if *ngroups_PCA*=2)

eigenvecd: k_{22} *x* k_{22} matrix of eigenvectors for dummy variables (if *ngroups_PCA*=2)

explvar: k_{21} *x 1* vector of explained variance (cumulative eigenvalues) for quantitative variables (if *ngroups_PCA*=2)

explvard: k_{22} *x 1* vector of explained variance (cumulative eigenvalues) for dummy variables (if *ngroups_PCA*=2)

nf: number of factors (if *ngroups_PCA*=2 then *nf* is 1 x 2 vector where the 1^{st} element corresponds to the number of factors for quantitative variables and the 2^{nd} element corresponds to the number of factors for dummy variables)

pc: *nobs x* k_{21} matrix of principal components (PC) (for quantitative variables (if *groups_PCA*=2))

pcd: *nobs* x k_{22} matrix of principal components (PC) for dummy variables (if *groups_PCA*=2)

WEIGHTS

AIC: 2^{k12} *x 1* vector of weights calculated based on the S-AIC weighting scheme

BIC: 2^{k12} *x 1* vector of weights calculated based on the S-BIC weighting scheme

equal: 2^{k12} *x 1* vector of weights calculated based on the equal weighting scheme

MMA: 2^{k12} *x 1* vector of weights calculated based on the MMA weighting scheme

incl_weights: *(3+flag_MMA) x* k_{12} matrix of inclusion weights. Ordering of rows corresponds to weighting schemes as: 1^{st} row - equal, 2^{nd} row - S-AIC, 3^{rd} row - S-BIC and 4^{th} row - MMA.

weights: 2^{k12} *x (3+flag_MMA)* matrix of *equal, AIC, BIC and MMA*