**Information about attached data and code**

## Data

**Abstract**

A set of risk factors that might contribute to the pricing of financial assets. The data are included in the attached R package and can be loaded in R via data(nfacdf)

**Availability**

Factors MKT, SMB, HML, RMW, CMA, MOM, QMJ, LIQ, BAB, and HMLD are publicly available.

Factors ME, ROE, and IA from the q-factor model are provided by Professor Lu Zhang. These three factors are not publicly available and should not be made publicly available without our permission.

**Description**

Factors MKT, SMB, HML, RMW, CMA, and MOM are publicly available at <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html>

Factor QMJ is publicly available at

<https://www.aqr.com/library/data-sets/quality-minus-junk-factors-monthly>

Factor LIQ is publicly available at

<http://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2016.txt>

Factor BAB is publicly available at

<https://www.aqr.com/library/data-sets/betting-against-beta-equity-factors-monthly>

Factor HMLD is publicly available at

<https://www.aqr.com/library/data-sets/the-devil-in-hmls-details-factors-monthly>

## Code

**Abstract**

The R package includes codes that implement the sampling the prior method, the sampling of posterior distribution, and the marginal likelihood computation.

**Description**

Main functions in the R package

* samplingthepriorre: sampling the prior method for restricted factor model
* samplingthepriorunre: sampling the prior method for unrestricted factor model
* MCMCfaclinearSDF: posterior sampling and marginal likelihood computation for Gaussian error factor model
* MCMCfaclinearSDFt: posterior sampling and marginal likelihood computation for Student-t error factor model

**Optional Information**

The R package depends on three other R packages, Rcpp, RcppArmadillo and coda.

## Instructions for Use

**Reproducibility**

* Reproduce the sampling the prior method in Table 4 and Table 5
  + Sampling the prior for Restricted model (Table 4). Take Model 1 (QMJ is the only x-factor) as an example.
    - Define two Boolean vectors. xind/wind stipulates columns in the data frame that go into x-factors/w-factors.

xind = c(T,F,F,F,F,T,T,T,F,F,F,F,F);

wind = c(F,T,T,T,T,F,F,F,T,T,T,T,T);

* Call the samplingthepriorre function

out = samplingthepriort(indx = xind,

indw = wind,

factordf = nfacdf, # data frame of risk factors

ntrain = 60, # first 60 obs in factordf are training sample

rho = 13+6, # prior rho parameter

slambda = .0025, # prior s\_lambda parameter

m = 100000, # number of simulated draws

nu = 2.1, # degrees of freedom for Student-t prior

nuf = 4, # degrees of freedom for Student-t model error

probs = c(.025,.975), # quantiles of simulated draws

seed = 100, # random seed

print.summary = TRUE)

* + Sampling the prior for unrestricted factor model (Table 5). The procedure is the same as in the case for restricted factor model except that now use the samplingthepriorunre function
* Reproduce the posterior sampling (Table 6) and marginal likelihood computation (Table 7). Take the q-factor model (MKT, ME, ROE, and IA are the only x-factors) as an example

outm = MCMCfaclinearSDFt(

xfrm = ~MKT+ME+ROE+IA, # x-factors

wfrm = ~SMB+HML+RMW+CMA+MOM+QMJ+LIQ+BAB+HMLD, # w-factors

data = nfacdf, # data frame of risk factors

ntrain = 60, # first 60 obs in data are training sample

rho = 13+6, # prior rho parameter

slambda = .0025, # prior s\_lambda parameter

n0 = 40000, # number of burn-in draws

n1 = 40000,

M = 10000, # number of MCMC draws

nu = 2.1, # degrees of freedom for Student-t prior

nuf = 4, # degrees of freedom for Student-t model error

seed = 123, # random seed

msg.print = TRUE);

**Replication (Optional)**

## Notes

Other relevant information, in particular how to access the data and code if not yet made publicly available.