Replication of a Real Estate Market Index

Bruno Remillard, et al.

2017 Industrial Problem Solving Workshop

August 7-11, 2017

Outline

- Introduction
- Our Approach
- Results
- Conclusion

The Problem

We are asked to replicate an index of illiquid assets (the real-estate market) using liquid tradeable securities.

$$dR_t^I = \left[\left(w_t^T R_t^f \right) + \left(1 - \sum_{i=1}^{d} w_{i,t} \right) r_t \right] dt + Q d\epsilon_t. \qquad (1)$$

- R_t^I is the returns of the index at time t,
- R_t^f is the vector of returns of tradeable securities at time t,
- w_t is a vector of $w_{i,t}$, containing the weight invested in each asset i at time t,
- r_t risk-free rate at time t (proxied by 1-month deposit rate),
- Q is the volatility of the replication errors.

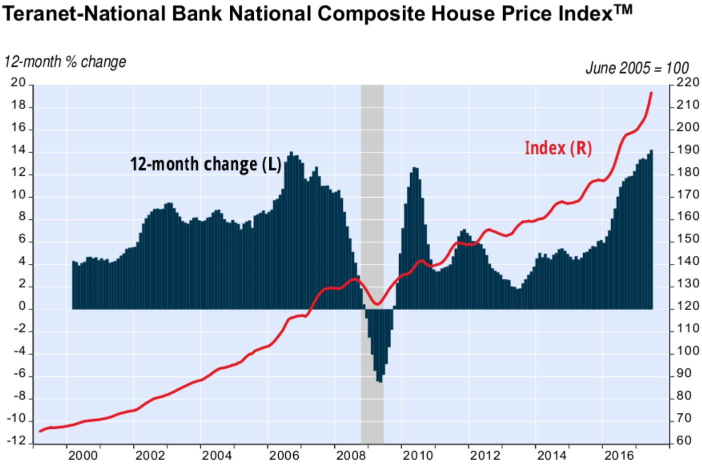
Question: why is this important? Because investors wish to be invested without buying actual houses.

Introduction - The Sketched Solution

In order to replicate the index, we need two key components:

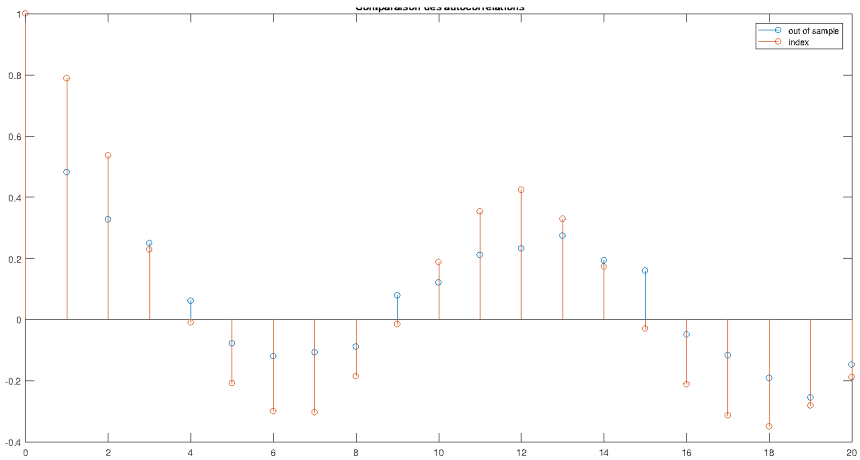
- The factors to include into the portfolio;
- The optimal weights w_t.

Introduction - Overview in Pictures - Index



Anything hidden from this picture? What happens within those 12 months...

Introduction - Overview in Pictures - Autocorrelation



Any chance of finding seasonality in our securities? Not in absence of arbitrage...

Introduction - Overview in Numbers

	Mean	Std	Skewness	Excess kurtosis
Index Return	6.31	2.23	-0.23	1.42

Table 1: Statistics of the index-return (1999-2017)

	Index.Return	CN.Comdty	CD.Curncy	CL.Comdty	GC.Comdty	HG.Comdty	ES.Index	HC.Index	PT.Index	BAA
Index.Return	1	0	0.110	0.100	0	0.110	0.080	0.010	0.080	0.170
CN.Comdty	0	1	-0.250	-0.240	0.200	-0.230	-0.230	-0.140	-0.180	0.050
CD.Curncy	0.110	-0.250	1	0.430	0.380	0.520	0.550	0.450	0.500	0.010
CL.Comdty	0.100	-0.240	0.430	1	0.210	0.380	0.200	0.290	0.370	-0.110
GC.Comdty	0	0.200	0.380	0.210	1	0.290	0.020	0.180	0.200	0.020
HG.Comdty	0.110	-0.230	0.520	0.380	0.290	1	0.450	0.450	0.480	0.020
ES.Index	0.080	-0.230	0.550	0.200	0.020	0.450	1	0.480	0.770	0.180
HC.Index	0.010	-0.140	0.450	0.290	0.180	0.450	0.480	1	0.460	-0.040
PT.Index	0.080	-0.180	0.500	0.370	0.200	0.480	0.770	0.460	1	0.050
BAA	0.170	0.050	0.010	-0.110	0.020	0.020	0.180	-0.040	0.050	1

Table 2: The correlation matrix for the data.

With the notable exception of the last security (BAA), which exhibits very high persistence in its returns, there is no significant autocorrelation of returns.

Our approach

- Factor selection
  - Adaptive Elastic-Net Regularization
  - Gradient boosting approach
- Weights Estimation
  - Rolling Regressions
  - Kalman filtering
  - Particle filtering

Factor Selection - Boosting

Extended Gradient-Boosting Learning

The goal of the learning algorithm in [2] is to identify the factors that significantly contribute to the model by minimizing the loss function,

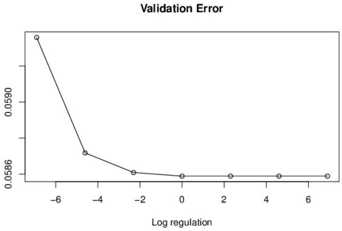
$$l(\bar{A}, \bar{w}, y, t, \lambda) = \sum_{t=1}^{d} (h_{\bar{w}}(\bar{A}_t) - y_t)^2 + \lambda |\bar{w}| \qquad (2)$$

where $h_{\bar{w}}(\bar{A}_t)$ is the value of the replicating portfolio at time t.

Factor Selection - Boosting

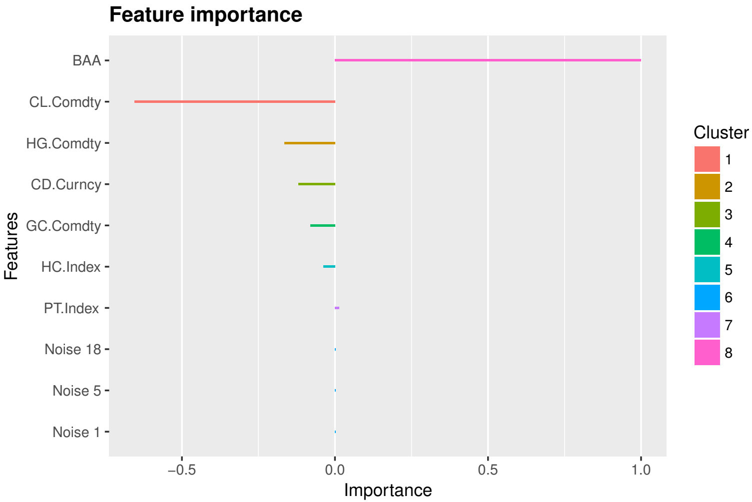
Gradient Boosting - Cross Validation

- We choose the most parsimonious learning-rate(λ) through Cross Validation:



Factor Selection - Results

- Using the selected λ we chose the following factors:



Factor Selection - Selected Factors

Table 3: Correlation matrix of the selected factors.

	Index.Return	CL.Comdty	HG.Comdty	HC.Index	BAA	BA.Comdty	TU.Comdty	FV.Comdty	TY.Comdty
Index.Return	1	0.092	0.113	0.012	0.208	0.260	-0.139	-0.152	-0.148
CL.Comdty	0.092	1	0.379	0.286	-0.112	0.182	-0.121	-0.147	-0.151
HG.Comdty	0.113	0.379	1	0.455	0.014	0.097	-0.179	-0.204	-0.210
HC.Index	0.012	0.286	0.455	1	-0.034	-0.030	-0.115	-0.110	-0.103
BAA	0.208	-0.112	0.014	-0.034	1	-0.042	-0.048	-0.029	-0.052
BA.Comdty	0.260	0.182	0.097	-0.030	-0.042	1	0.017	-0.057	-0.078
TU.Comdty	-0.139	-0.121	-0.179	-0.115	-0.048	0.017	1	0.874	0.766
FV.Comdty	-0.152	-0.147	-0.204	-0.110	-0.029	-0.057	0.874	1	0.961
TY.Comdty	-0.148	-0.151	-0.210	-0.103	-0.052	-0.078	0.766	0.961	1

Factor Selection - Adaptive Elastic-Net Regularization

Adaptive Elastic-Net

Performs variable selection for linear models with constant weights, and estimates the weights $\hat{w}$ as

$$\hat{w} \triangleq \left(1 + \frac{\alpha}{d}\right) \left[\operatorname{argmin}_{w \in \mathbb{R}^d} \left(\sum_{j=1}^{d} (h_{\bar{w}}(\bar{A}_t) - y_t)^2 + \lambda \sum_{j=1}^{d} \hat{\eta}_j |\bar{w}| + \alpha (\bar{w})^2 \right)\right]$$

$$\hat{\eta}_j \triangleq (|\hat{w}_j(Enet)|)^{-\gamma},$$

where $\hat{w}_j(Enet)$ are the optimal weights defined in [9]. The Adaptive Elastic-Net algorithm was shown in [10] to converge to the true set of explanatory variables under certain assumptions.

Factor Selection - Adaptive Elastic-Net Regularization

Adaptive Elastic-Net

Performs variable selection for linear models with constant weights, and estimates the weights $\hat{w}$ as

$$\hat{w} \triangleq \left(1+\frac{\alpha}{d}\right)\left[\operatorname{argmin}_{w\in\mathbb{R}^d}\left(\sum_{j=1}^{d}(h_{\bar{w}}(\bar{A}_t)-y_t)^2+\lambda\sum_{j=1}^{d}\hat{\eta}_j|\bar{w}|+\alpha(\bar{w})^2\right)\right]$$

$$\hat{\eta}_j \triangleq (|\hat{w}_j(Enet)|)^{-\gamma}\,,$$

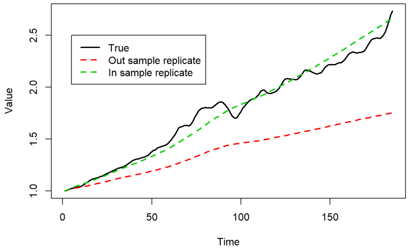
where $\hat{w}_j(Enet)$ are the optimal weights defined in [9]. The Adaptive Elastic-Net algorithm was shown in [10] to converge to the true set of explanatory variables under certain assumptions.

- The adaptive elastic-net was used to perform factor selection.
- Upon cross-validation, all factors were considered as important.

Estimating Weights - Rolling Regressions

Result

We performed a 24 month moving window and here is the result:



Conclusion

- ✓ Provides a benchmark and starting point for the filter,
- × The regression does not provide an accurate forecast.

Estimating Weights - Filtering

Description of the Filtering Problem

- We are interested in predicting an unobservable *signal* $\bar{X}_t$,
- However we only know the *observation* process Y_t which depends on $\bar{X}_t$ but is obscured by noise.
- The goal is to best guess the conditional density of $\bar{X}_t$ given Y_t, denoted by $\pi_t(f)$ (see [4] for details).

Estimating Weights - Filtering

Description of the Filtering Problem

- We are interested in predicting an unobservable *signal* $\bar{X}_t$,
- However we only know the *observation* process Y_t which depends on $\bar{X}_t$ but is obscured by noise.
- The goal is to best guess the conditional density of $\bar{X}_t$ given Y_t, denoted by $\pi_t(f)$ (see [4] for details).

Kalman-Bucy Filter

✓ Simple closed form,

× Relies on idealized assumptions.

Estimating Weights - Filtering

Description of the Filtering Problem

- We are interested in predicting an unobservable *signal* $\bar{X}_t$,
- However we only know the *observation* process Y_t which depends on $\bar{X}_t$ but is obscured by noise.
- The goal is to best guess the conditional density of $\bar{X}_t$ given Y_t, denoted by $\pi_t(f)$ (see [4] for details).

Kalman-Bucy Filter

✓ Simple closed form,

× Relies on idealized assumptions.

Particle Filter

✓ Difficult to calibrate initial state,

× Does not have a simple closed-form expression.

Estimating Weights - Filtering - Methodologies

- Initialize Kalman Filter with the regression weights;
- Use the Kalman estimates of the model parameters for the Particle Filter.

Estimating Weights - Filtering - Model

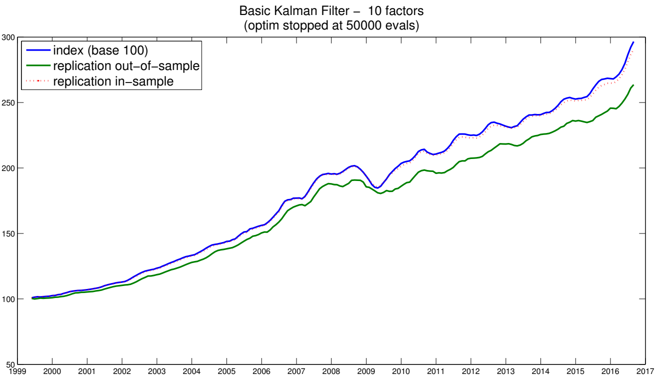
We use the following model:

$$\begin{aligned} W_t &= \phi W_{t-1} + \epsilon_t \\ R_t^{I*} &= W_t^T R_t^{f*} + \xi_t \end{aligned} \tag{3}$$

- W_t, the portfolio weights' matrix, is the signal being filtered;
- ϕ is a constant diagonal (for now) matrix capturing dynamics of the weights;
- R_t^{I*} is the observation, in our case the excess return of the index;
- R_t^{f*} are the excess returns of our selected factors;
- ϵ and ξ are IID and Normally distributed.

Estimating Weights - Filtering - Kalman

Using selected factors from the Gradient-Boosting we get:

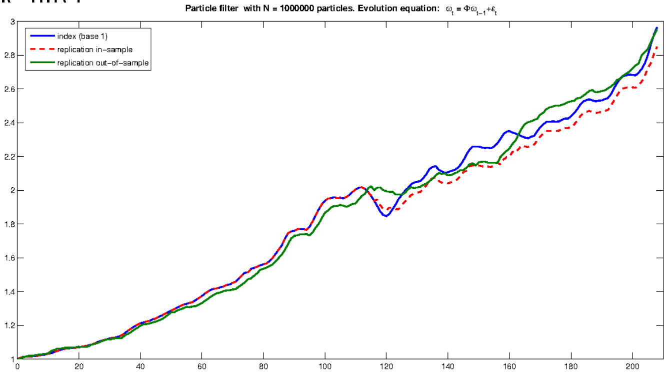


We tried to improve this method by:

- Model 2: including the original scope (14 factors, minus 3 highly correlated / redundant assets);
- Model 3: initializing the filter to weights which were estimated using the regression approach on the first 36 months.

Estimating Weights - Filtering - Particles

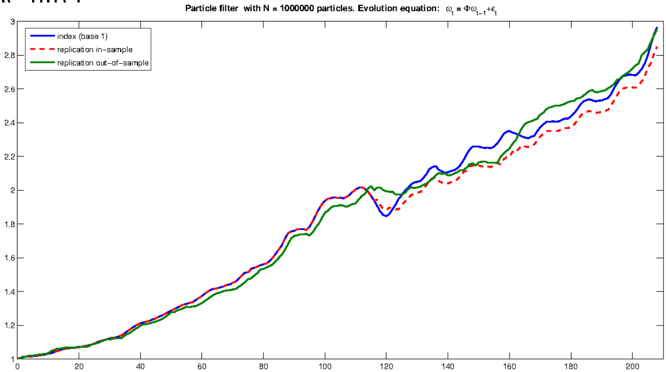
We obtained a further improvement out-of-sample using the particle filter:



Problem solved?

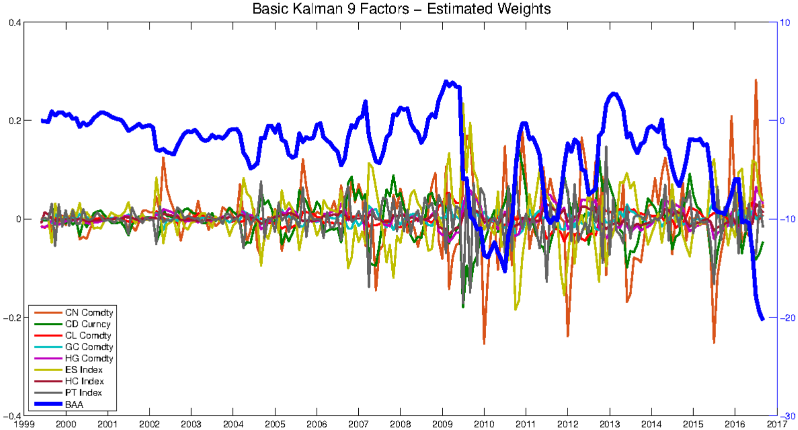
Estimating Weights - Filtering - Particles

We obtained a further improvement out-of-sample using the particle filter:



Problem solved? Not quite...

Results - Behind the Scenes



Although most weights vary within acceptable boundaries:

- Some have a negative value for ϕ, implying alternation of long/short positions $\implies$ costly!
- One security is leveraged up to 20 times $\implies$ risky!

Results - Statistics

Portfolio	TE	Pearson Corr	Kendall Corr	Mean	Std	Skew	Excess kurt
Target	0	1	1	6.3130	2.2388	-0.2256	1.4168
Regression	6.0823	0.1331	0.0472	0.0053	0.0011	0.5543	2.0818
Kalman 1	1.3391	0.9875	0.8918	6.1752	2.0605	-0.1797	1.5414
Kalman 2	6.74e-10	1	1	6.2954	2.2430	-0.2191	1.4029
Kalman 3	7.84e-10	0.5763	0.3946	6.2738	1.7355	- 0.2416	1.0534
Particle (Unc.)	4.2000	0.8450	0.8568	6.0789	2.0887	-0.1913	1.4587
Particle (Con.)	3.3496	0.9056	0.9270	5.7817	2.1457	-0.3291	1.5756

Table 4: In-sample statistics.

Portfolio	TE	Pearson Corr	Kendall Corr	Mean	Std	Skew	Excess kurt
Target	0	1	1	6.3130	2.2388	-0.2256	1.4168
Regression	9.8850	0.1284	0.0132	0.0031	0.0011	0.5682	2.1050
Kalman 1	6.0297	0.6587	0.4420	5.6274	1.8757	-0.1584	1.7168
Kalman 2	6.5171	0.5763	0.3946	6.2738	2.2430	1.7355	1.0534
Kalman 3	6.5033	0.5777	0.3937	6.2792	1.7307	0.2609	1.0423
Particle (Unc.)	8.0168	0.3685	0.2484	6.2779	1.8499	0.2040	0.4957
Particle (Con.)	7.2842	0.5038	0.3306	6.6414	1.9172	0.0092	1.2384

Table 5: Out-of-sample statistics.

Conclusion - Extensions and Further Research

- Current extensions:
  - Estimation of ϕ using a constrained optimization in the Kalman Filter;
  - The results on the weights is significant (leverage on BAA drops to -6) but still unsatisfactory.
- Further research:
  - Restricting weights-space using the Particle Filter (easy to implement but very computer intensive);
  - Integrating seasonality using the following model:

$$W_t = \phi W_{t-1} + \gamma W_{t-12} + \epsilon_t \tag{4}$$

  - Studying a larger scope and reducing dimensionality (as well as rebalancing costs!) using Boosting.

References

[1] A. Bain and D. Crisan. *Fundamentals of stochastic filtering*, volume 60 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2009.

[2] P. Bühlmann and T. Hothorn. Boosting algorithms: regularization, prediction and model fitting. *Statist. Sci.*, 22(4):477–505, 2007.

[3] R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Trans. ASME Ser. D. J. Basic Engrg.*, 83:95–108, 1961.

[4] R. S. Liptser and A. N. Shiryaev. *Statistics of random processes. I*, volume 5 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, expanded edition, 2001. General theory, Translated from the 1974 Russian original by A. B. Aries, Stochastic Modelling and Applied Probability.

[5] M. K. Pitt. Smooth particle filters for likelihood evaluation and maximisation. 2002.

[6] B. Rémillard. *Statistical methods for financial engineering*. CRC Press, Boca Raton, FL, 2013.

[7] B. Remillard, B. Nasri, and M. Ben-abdellatif. Replication methods for financial indexes. 2017.

[8] T. Zhang and B. Yu. Boosting with early stopping: convergence and consistency. *Ann. Statist.*, 33(4):1538–1579, 2005.

[9] H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 67(2):301–320, 2005.

[10] H. Zou and H. H. Zhang. On the adaptive elastic-net with a diverging number of parameters. *Ann. Statist.*, 37(4):1733–1751, 2009.

Acknowledgements

Thank You!