Identifying Small Mean Reverting Portfolios

By Alexandre d'Aspremont*

August 22, 2007

Abstract

Given asset price time series, we study the problem of forming portfolios with maximum mean reversion while constraining the number of assets in these portfolios. We show that this problem is equivalent to sparse canonical correlation analysis and study various algorithms to solve the corresponding sparse generalized eigenvalue problems. Finally, we test the performance of various convergence trading strategies on these portfolios.

1 Introduction

Mean reversion has received a very significant amount of attention as a classic indicator of predictability in financial markets and is often apparent in equity index returns over long horizons. While mean reversion is easy to identify in univariate time series, isolating portfolios of asset with maximum mean reversion is a much more complex problem. Classic solutions include cointegration or canonical correlation analysis, which will be discussed in what follows.

The problem with theses techniques though is that the mean reverting (or momentum) portfolios they identify are often dense, i.e. they include every asset in the time series analyzed. For arbitrageurs, this is an issue because exploiting the corresponding statistical arbitrage opportunities involves considerable *transaction costs*. For econometricians, this is also a problem since it impacts the *interpretability* of the resulting portfolio. Finally, optimally mean reverting portfolios often behave like noise and sometimes vary well inside bid-ask spreads, hence do not form meaningful statistical arbitrage opportunities.

Here, we would like to argue that seeking *sparse* portfolios instead, i.e. optimally mean reverting portfolios with a few assets, solves many of these issues at once. Fewer assets means less transaction costs, and more interpretable results. Although we will show that the tradeoff between mean reversion and sparsity is often very favorable, penalizing for sparsity also means that sparse portfolios vary in a wider range of prices, which means that the market inefficiencies it isolates are more significant.

Mean reversion has of course received a considerable amount of attention in the literature, most authors, such as Fama & French (1988), Poterba & Summers (1988) among many others, using it

*ORFE, Princeton University, Princeton, NJ 08544, USA. Email: aspremon@princeton.edu

to model and test for market predictability in excess returns. Cointegration techniques, (see Engle & Granger (1987), and Alexander (1999) for a survey of applications in finance) are usually used to extract mean reverting portfolios from multivariate time series. Early techniques relied on a mix of regression and Dickey & Fuller (1979) stationarity tests or Johansen (1988) type tests. It was subsequently discovered that an earlier decomposition technique due to Box & Tiao (1977) could be used to extract cointegrated vectors by solving a generalized eigenvalue problem, see Bewley, Orden, Yang & Fisher (1994) for a discussion. Several authors then focused on the optimal investment problem when expected returns are mean reverting, with Kim & Omberg (1996) and Campbell & Viceira (1999) or Wachter (2002) for example obtaining closed-form solutions in some particular cases. Liu & Longstaff (2004) study the optimal investment problem in the presence of a "textbook" finite horizon arbitrage opportunity, modeled as a Brownian bridge, while Jurek & Yang (2006) study this same problem when the arbitrage horizon is indeterminate. Gatev, Goetzmann & Rouwenhorst (2006) also studied the performance of pairs trading, which are classic examples of structurally mean-reverting portfolios. Finally, the LTCM meltdown in 1998 focused a lot of attention on the impact of leverage limits and liquidity, see Grossman & Vila (1992) or Xiong (2001) for a discussion.

Our contribution here is twofold. First, we describe two algorithms for extracting sparse mean reverting portfolios from multivariate time series. One is based on a simple greedy search on the list of assets to include. The other uses semidefinite relaxation techniques to directly get good solutions. Second, we study the sparsity mean reversion tradeoff in several markets, then test the impact of these portfolios' predictability using various convergence trading strategies.

The paper is organized as follows. In Section 2, we detail two algorithms to extract small mean reverting portfolios from multivariate data sets. In Section 3, we show how to form optimal portfolios of assets following an Ornstein-Uhlenbeck process. Finally, we present some empirical results in Section 4.

2 Small mean reverting portfolios

Let S_{ti} be the value at time t of an asset S_i for $i = 1, \ldots, n$ and $t = 1, \ldots, m$. In what follows, our objective will be to form portfolios P_t of assets with coefficients x_i which follow an Ornstein-Uhlenbeck process given by:

$$dP_t = \lambda(\bar{P} - P_t)dt + \sigma dZ_t \quad \text{with } P_t = \sum_{i=1}^{n} x_i S_{ti}$$

where Z_t is a standard Brownian motion. We will seek to maximize the mean reversion coefficient λ of P_t by adjusting the coefficients x_i, under the constraints that $\|x\| = 1$ (for) and that the cardinality of x, *i.e.* the number of nonzero coefficients in x, remains below a certain level $k > 0$.

2.1 Canonical decompositions

For time series analysis in this section, we will work in a discrete setting and assume that the asset prices follow a (stationary) autoregressive process with:

$$S_t = AS_{t-1} + Z_t \tag{1}$$

where S_{t-1} is the lagged portfolio process, $A \in \mathbf{R}^{n\times n}$ and Z_t is a vector of i.i.d. Gaussian noise with zero mean and covariance $\Sigma \in \mathbf{S}^n$, independent of S_{t-1}. The canonical analysis in Box & Tiao (1977) starts as follows. Suppose we take $n = 1$ in equation (1) above, we get:

$$\mathbf{E}[S_t^2] = \mathbf{E}[(AS_{t-1})^2] + \mathbf{E}\left[Z_t^2\right]$$

which can be rewritten as $\sigma_t^2 = \sigma_{t-1}^2 + \Sigma$. Box & Tiao (1977) then measure the predictability of stationary series by:

$$\lambda = \frac{\sigma_{t-1}^2}{\sigma_t^2}. \tag{2}$$

Suppose now that we consider a portfolio $P_t = x^T S_t$ with $x \in \mathbf{R}^n$, using (1) we know that $x^T S_t = x^T AS_{t-1} + x^T Z_t$, so its predicability can be measured as:

$$\lambda_x = \frac{x^T A\Gamma A^T x}{x^T \Gamma x}$$

where $\Gamma = \mathbf{E}[SS^T]$. This means that the portfolio with maximum (resp. minimum) predictability will be the eigenvector corresponding to the largest (resp. smallest) eigenvalue of the the matrix:

$$\Gamma^{-1} A\Gamma A^T. \tag{3}$$

Bewley et al. (1994) show that both the Box-Tiao canonical decompositions and the Johansen maximum likelihood decompositions can be be formulated in this manner and we briefly recall their result below. Following Bewley et al. (1994), equation (1) can be rewritten:

$$S_t = \hat{S}_t + \hat{Z}_t$$

where $\hat{S}_t$ is the least squares estimate of S_t with $\hat{S}_t = S_{t-1}\hat{A}$ and:

$$\hat{A} = \left(S_{t-1}^T S_{t-1}\right)^{-1} S_{t-1}^T S_t.$$

Box-Tiao procedure. The Box & Tiao (1977) procedure then finds linear combinations of the assets ranked in order of predictability by computing the eigenvectors of the matrix:

$$\left(S^T S\right)^{-1} \left(\hat{S}_t^T \hat{S}_t\right) \tag{4}$$

where is $\hat{S}_t$ is the least squares estimate computed above.

Johansen procedure. Following Bewley et al. (1994), the maximum likelihood procedure for estimating cointegrating vectors derived in Johansen (1991) can also be written as a canonical decomposition à la Box & Tiao (1977). This time however, we perform a canonical analysis between the first order differences of the series S_t and their lagged values S_{t-1}, so we rewrite equation (1) as:

$$\Delta S_t = Q S_{t-1} + Z_t$$

where $Q = A - \mathbf{I}$. The basis of (potentially) cointegrating portfolios is then found by solving the following generalized eigenvalue problem:

$$\lambda S_{t-1}^T S_{t-1} - S_{t-1}^T \Delta S_t (\Delta S_t^T \Delta S_t)^{-1} \Delta S_t^T S_{t-1} \tag{5}$$

in the variable $\lambda \in \mathbf{R}$.

2.2 Sparse generalized eigenvalue problems

Both problems above can be written as generalized eigenvalue problems of the form:

$$\det(\lambda A - B) = 0 \tag{6}$$

in the variable $\lambda \in \mathbf{R}$, where $A, B \in \mathbf{S}^n$. This is usually solved using a QZ decomposition. The largest solution of this problem can be written in variational form as:

$$\lambda^{\max} = \max_{x \in \mathbf{R}^n} \frac{x^T B x}{x^T A x}.$$

Here however, we seek to maximize that ratio while constraining the cardinality of the (portfolio) coefficient vector x and solve instead:

$$\begin{array}{ll} \text{maximize} & x^T A x / x^T B x \\ \text{subject to} & \mathbf{Card}(x) \leq k \\ & \|x\| = 1, \end{array} \tag{7}$$

where $k > 0$ is a given constant and $\mathbf{Card}(x)$ is the number of nonzero coefficients in x. This will compute a sparse portfolio with maximum predictability (or momentum), a similar problem can be formed to minimize it (and obtain small mean reverting portfolios). This is a (hard) combinatorial problem, so we can't expect to get optimal solutions and we discuss below two efficient techniques to get good approximate solutions.

Greedy optimization. Let us call I the support of the vector x in problem (7) above:

$$I_k = \{i \in [1, n] : \ x_i \neq 0\},$$

so by construction $|I_k| \leq k$. Let us now build approximate solutions to (7) recursively in k. When $k = 1$, we can simply find I_1 as:

$$I_1 = \underset{i \in [1,n]}{\operatorname{argmax}}\, A_{ii} / B_{ii}.$$

Suppose now that we have a good approximate solution with support set I_k given by:

$$x_k = \underset{\{x \in \mathbf{R}^n:\, x_{I_k^c}=0\}}{\operatorname{argmax}} \frac{x^T B x}{x^T A x},$$

which can be solved as a generalized eigenvalue problem of size k. We seek to add one variable with index i_{k+1} to the set I_k to produce the largest increase in predictability. The index i_{k+1} can be computed as:

$$i_{k+1} = \underset{i}{\operatorname{argmax}} \max_{\{x \in \mathbf{R}^n:\, x_{J_i^c}=0\}} \frac{x^T B x}{x^T A x}, \quad \text{where } J_i = I_k \cup \{i\}$$

and we can then define $I_{k+1} = I_k \cup \{i_{k+1}\}$. Naturally, the optimal solutions of problem (7) might not have increasing support sets $I_k \subset I_{k+1}$ hence the solutions found by this recursive algorithm are potentially far from optimal. However, the cost of this method is relatively low: with each iteration costing $O(k^2(n-k))$, the complexity of computing solutions for all k is in $O(n^4)$. This recursive procedure can also be repeated both forward and backward to improve the quality of the solution.

Semidefinite relaxation. Here, we derive a semidefinite relaxation for sparse generalized eigenvalue problems in (7):

$$\begin{array}{ll} \text{maximize} & x^T A x / x^T B x \\ \text{subject to} & \mathbf{Card}(x) \leq k \\ & \|x\| = 1, \end{array}$$

with variable $x \in \mathbf{R}^n$. As in d'Aspremont, El Ghaoui, Jordan & Lanckriet (2007), we can form an equivalent program in terms of $X = xx^T \in \mathbf{S}_n$:

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(AX)/\mathbf{Tr}(BX) \\ \text{subject to} & \mathbf{Card}(X) \leq k^2 \\ & \mathbf{Tr}(X) = 1 \\ & X \succeq 0, \ \mathbf{Rank}(X) = 1, \end{array}$$

in the variable $X \in \mathbf{S}_n$. This program is equivalent to the first one: if X is a solution to the above problem, then $X \succeq 0$ and $\mathbf{Rank}(X) = 1$ mean that we have $X = xx^T$, while $\mathbf{Tr}(X) = 1$ implies that $\|x\|_2 = 1$. Finally, if $X = xx^T$ then $\mathbf{Card}(X) \leq k^2$ is equivalent to $\mathbf{Card}(x) \leq k$. Now, since $\mathbf{Card}(u) = q$ implies $\|u\|_1 \leq \sqrt{q}\|u\|_2$, we can replace the nonconvex constraint $\mathbf{Card}(X) \leq k^2$, by a weaker but convex constraint: $\mathbf{1}^T|X|\mathbf{1} \leq k$, using $\|X\|_F = \sqrt{x^T x} = 1$ when $X = xx^T$ and $\mathbf{Tr}(X) = 1$. We can then relax (in fact, drop) the rank constraint in this program to get the following quasi-convex program:

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(AX)/\mathbf{Tr}(BX) \\ \text{subject to} & \mathbf{1}^T|X|\mathbf{1} \leq k \\ & \mathbf{Tr}(X) = 1 \\ & X \succeq 0, \end{array} \tag{8}$$

in the variable $X \in \mathbf{S}_n$. We do the following change of variables:

$$Y = \frac{X}{\mathbf{Tr}(BX)}, \quad z = \frac{1}{\mathbf{Tr}(BX)}$$

and solve:

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(AY) \\ \text{subject to} & \mathbf{1}^T|Y|\mathbf{1} - kz \leq 0 \\ & \mathbf{Tr}(Y) - z = 0 \\ & \mathbf{Tr}(BY) = 1 \\ & Y \succeq 0, \end{array} \tag{9}$$

which is a semidefinite program (SDP) in the variables $Y \in \mathbf{S}_n$ and $z \in \mathbf{R}_+$ and can be solved using standard SDP solvers such as SEDUMI by Sturm (1999) and SDPT3 by Toh, Todd & Tutuncu (1999). We can extract an approximate solution to (7) as the rescaled dominant eigenvector of the optimal solution to (9).

3 Dynamic portfolio selection

In the previous section, we have showed how to extract small mean reverting (or predictable) portfolios from asset time series. In this section we assume that we have identified such a mean reverting portfolio with dynamics given by:

$$dP_t = \lambda(\bar{P} - P_t)dt + \sigma dZ_t, \tag{10}$$

and we detail how to optimally trade these portfolios under various assumptions regarding market friction and risk-management constraints. We begin by quickly recalling results on estimating Ornstein-Uhlenbeck processes.

3.1 Estimating Ornstein-Uhlenbeck processes

By integrating P_t over a time increment Δt we get:

$$P_t = \bar{P} + e^{-\lambda \Delta t}(P_{t-\Delta t} - \bar{P}) + \sigma \int_{t-\Delta t}^{t} e^{\lambda(s-t)} dZ_s,$$

which means that we can estimate λ and σ by simply regressing P_t on $P_{t-\Delta t}$ and a constant. With

$$\int_{t-\Delta t}^{t} e^{\lambda(s-t)} dZ_s \sim \sqrt{\frac{1 - e^{-2\lambda \Delta t}}{2\lambda}}\, \mathcal{N}(0,1),$$

we get the following estimators for the parameters of P_t:

$$\hat{\mu} = \frac{1}{N} \sum_{i=0}^{N} P_{t_i}$$

$$\begin{aligned}\hat{\lambda} &= -\frac{1}{\Delta t}\log\left(\frac{\sum_{i=1}^{N}(P_{t_i}-\hat{\mu})(P_{t_{i-1}}-\hat{\mu})}{\sum_{i=1}^{N}(P_{t_i}-\hat{\mu})(P_{t_i}-\hat{\mu})}\right)\\ \hat{\sigma} &= \sqrt{\frac{2\lambda}{(1-e^{-2\lambda\Delta t})(N-2)}\sum_{i=1}^{N}\left((P_{t_i}-\hat{\mu})-e^{-\lambda\Delta t}(P_{t_{i-1}}-\hat{\mu})\right)^2}\end{aligned}$$

where $t_i = i\Delta t$, for $i = 0, \ldots, N$ and $t_N = T$.

3.2 Utility maximization in frictionless markets

Suppose now that an agent invests in an asset P_t and in a riskless bond B_t following:

$$dB_t = rB_t dt,$$

the wealth W_t of this agent will follow:

$$dW_t = N_t dP_t + (W_t - N_t P_t) r dt.$$

If P_t follows a mean reverting process given by (10), this is also:

$$dW_t = (r(W_t - N_t P_t) + \lambda(\bar{P} - P_t)N_t)dt + N_t \sigma dZ_t.$$

If we write the value function:

$$V(W_t, P_t, t) = \max_{N_\bullet} \mathbf{E}_t\left[e^{-\beta(T-t)}U(W_t)\right],$$

the H.J.B. equation for this problem can be written:

$$\begin{aligned}\beta V &= \max_{N_t} \frac{\partial V}{\partial P}\lambda(\bar{P}_t - P_t) + \frac{\partial V}{\partial W}(r(W_t - N_t P_t) + \lambda(\bar{P} - P_t)N_t) + \frac{\partial V}{\partial t}\\ &\quad + \frac{1}{2}\frac{\partial^2 V}{\partial P^2}\sigma^2 + \frac{1}{2}\frac{\partial^2 V}{\partial P \partial W}N_t\sigma^2 + \frac{1}{2}\frac{\partial^2 V}{\partial W^2}N_t^2\sigma^2\end{aligned}$$

Maximizing in N_t yields the following expression for the number of shares in the optimal portfolio:

$$N_t = \frac{\partial V/\partial W}{\partial^2 V/\partial W^2 \sigma^2}(\lambda(\bar{P} - P_t) - rP_t) - \frac{\partial^2 V/\partial P \partial W}{\partial^2 V/\partial W^2} \tag{11}$$

Jurek & Yang (2006) solve this equation explicitly for $U(x) = \log x$ and $U(x) = x^{1-\gamma}/(1-\gamma)$ and we recover in particular the classic expression:

$$N_t = \left(\frac{\lambda(\bar{P} - P_t) - rP_t}{\sigma^2}\right)W_t,$$

in the log-utility case.

3.3 Leverage constraints

Suppose now that the portfolio is subject to fund withdrawals so that the total wealth evolves according to:

$$dW = d\Pi + dF$$

where $d\Pi = N_t dP_t + (W_t - N_t P_t) r dt$ and dF represents fund flows, with:

$$dF = f d\Pi + \sigma_f dZ_t^{(2)}$$

where $Z_t^{(2)}$ is a Brownian motion (independent of Z_t). Jurek & Yang (2006) show that the optimal portfolio allocation can also be computed explicitly in the presence of fund flows, with:

$$N_t = \left(\frac{\lambda(\bar{P} - P_t) - r P_t}{\sigma^2} \right) \frac{1}{(1+f)} W_t = L_t W_t,$$

in the log-utility case.

The constant f also implicitly limits leverage. In steady state, we have:

$$P_t \sim \mathcal{N}\left(\bar{P}, \frac{\sigma^2}{2\lambda}\right)$$

which means that the leverage L_t itself is normally distributed. If we assume for simplicity that $\bar{P} = 0$, given the fund flow parameter f, the leverage will remain below the level M given by:

$$M = \frac{\alpha(\lambda + r)}{(1+f)\sigma\sqrt{2\lambda}}$$

with confidence level $N(\alpha)$, where $N(x)$ is the Gaussian CDF. The bound on leverage M can thus be seen as an alternate way of identifying or specifying the fund flow constant f in order to manage capital outflow risks.

4 Empirical results

In this section, we first compare the performance of the algorithms described in Section 2. We then study the mean reversion versus sparsity tradeoff on various financial instruments. Finally, we test the performance of the convergence trading strategies described in Section 3.

4.1 Numerical performance

In Figure 1 we plot the result of the Box-Tiao decomposition on U.S. swap rate data (see details below). Each portfolio is a *dense* linear combination of swap rates, ranked in decreasing order of predictability. In Figure 2, we apply the greedy search algorithm detailed in Section 2.2 to the same data set, we plot the *sparse* portfolio processes obtained by constraining the sparsity of the portfolios. Each subplot lists the number k of nonzero coefficients of the corresponding portfolio and its mean reversion coefficient λ. Finally, Figure 3 compares the performance of the greedy search algorithm versus the semidefinite relaxation derived in Section 2.2. For each algorithm, we plot the mean reversion coefficient λ versus portfolio cardinality (number of nonzero coefficients). We observe on this example that while the semidefinite relaxation does produce better results in some instances, the greedy search is more reliable. Of course, both algorithms recover the same solutions when the target cardinality is set to $k = 1$ or $k = n$.

4.2 Mean reversion versus sparsity

In this section, we study the mean reversion versus sparsity tradeoff on several data sets. We also test the persistence of this mean reversion out of sample.

Swap rates. In Figure 4 we compare in and out of sample estimates of the mean reversion versus cardinality tradeoff. We study U.S. swap rate data for maturities 1Y, 2Y, 3Y, 4Y, 5Y, 7Y, 10Y and 30Y from 1998 until 2005. We first use the greedy algorithm of Section 2.2 to compute optimally mean reverting portfolios of increasing cardinality for time windows of 200 days and repeat the procedure every 50 days. We plot average mean reversion versus cardinality in Figure 4 on the left. We then repeat the procedure, this time computing the mean reversion in the 200 days (out of sample) time window immediately following our sample and also plot average mean reversion versus cardinality. In Figure 4 on the right, we plot the out of sample portfolio price range (spread between min. and max. in basis points) versus cardinality (number of nonzero coefficients) on the same U.S. swap rate data.

4.3 Optimal portfolios

Here, we measure the performance of the convergence trading strategies detailed in Section 3. In Figure 5 we plot average out of sample sharpe ratio versus portfolio cardinality on a 50 days (out of sample) time window immediately following the 100 days over which we estimate the process parameters.

References

Alexander, C. (1999), ‘Optimal hedging using cointegration’, *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* **357**(1758), 2039–2058.

Bewley, R., Orden, D., Yang, M. & Fisher, L. (1994), ‘Comparison of Box-Tiao and Johansen Canonical Estimators of Cointegrating Vectors in VEC (1) Models’, *Journal of Econometrics* **64**, 3–27.

Box, G. E. & Tiao, G. C. (1977), ‘A canonical analysis of multiple time series’, *Biometrika* **64**(2), 355.

Campbell, J. & Viceira, L. (1999), ‘Consumption and Portfolio Decisions When Expected Returns Are Time Varying’, *The Quarterly Journal of Economics* **114**(2), 433–495.

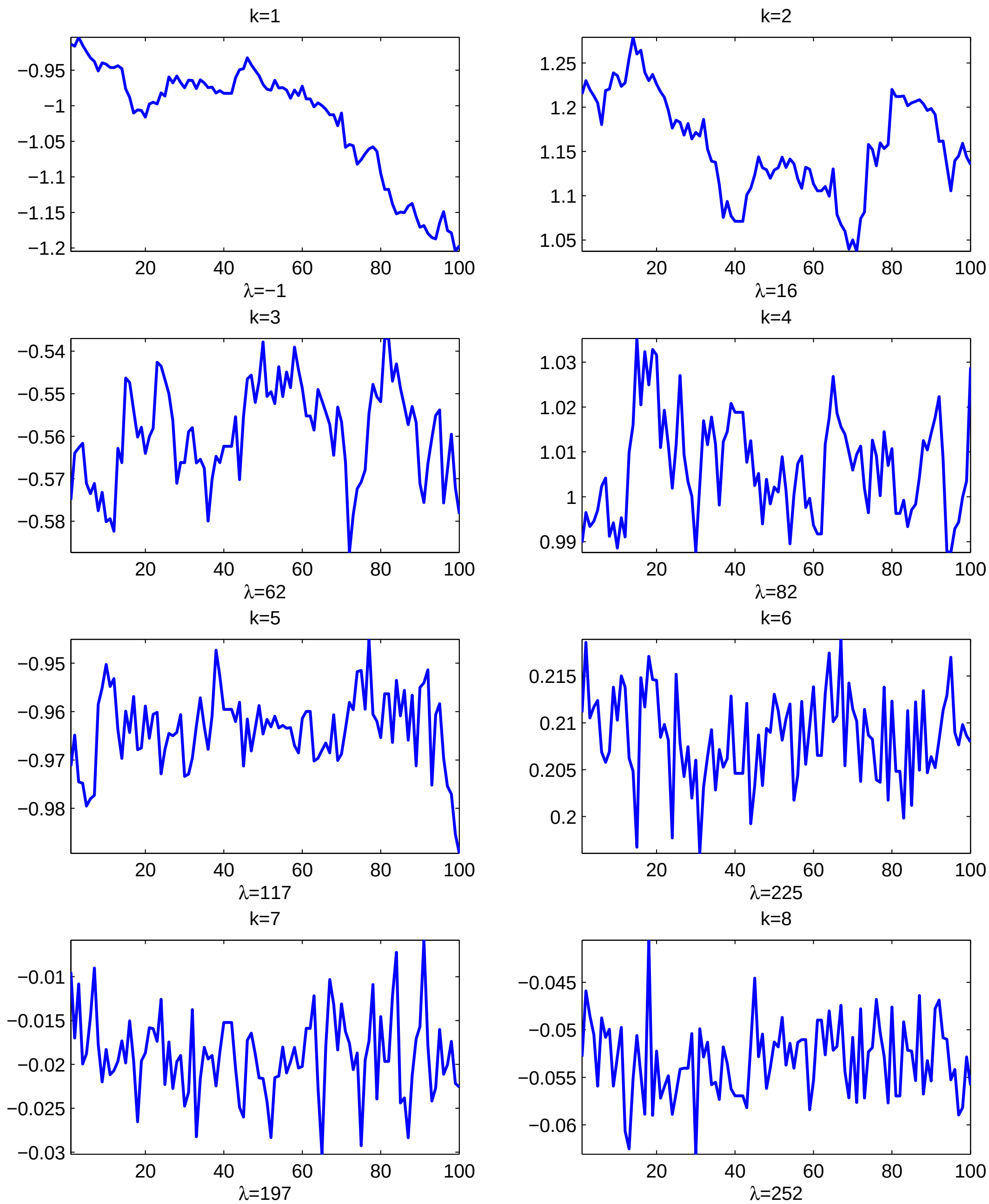

Figure 1: Box-Tiao decomposition on 100 days of U.S. swap rate data (in percent), ranked in decreasing order of predictability. The mean reversion coefficient λ is listed below each plot.

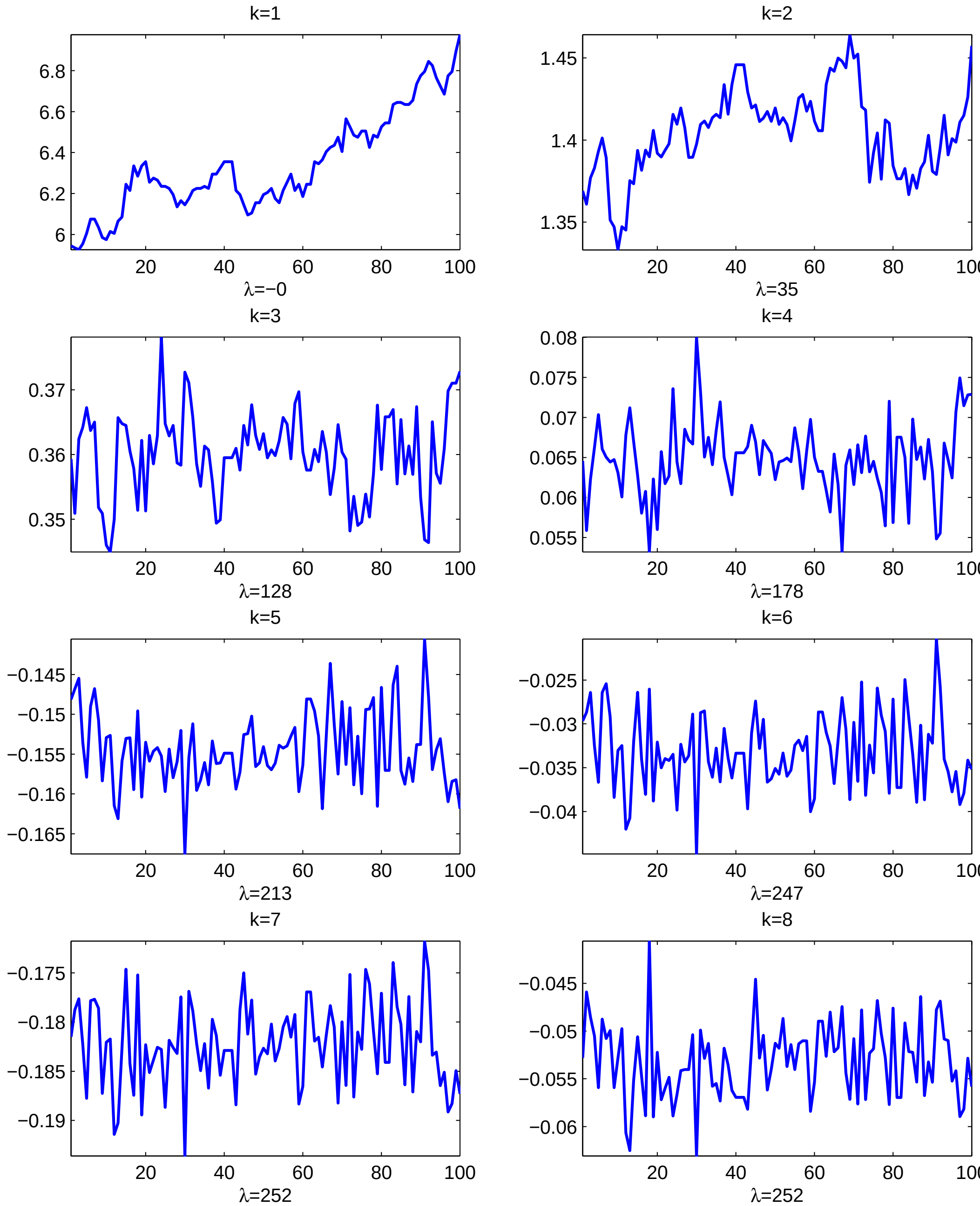

Figure 2: Sparse canonical decomposition on 100 days of U.S. swap rate data (in percent). The number of nonzero coefficients in each portfolio vector is listed as k on top of each subplot, while the mean reversion coefficient λ is listed below each one.

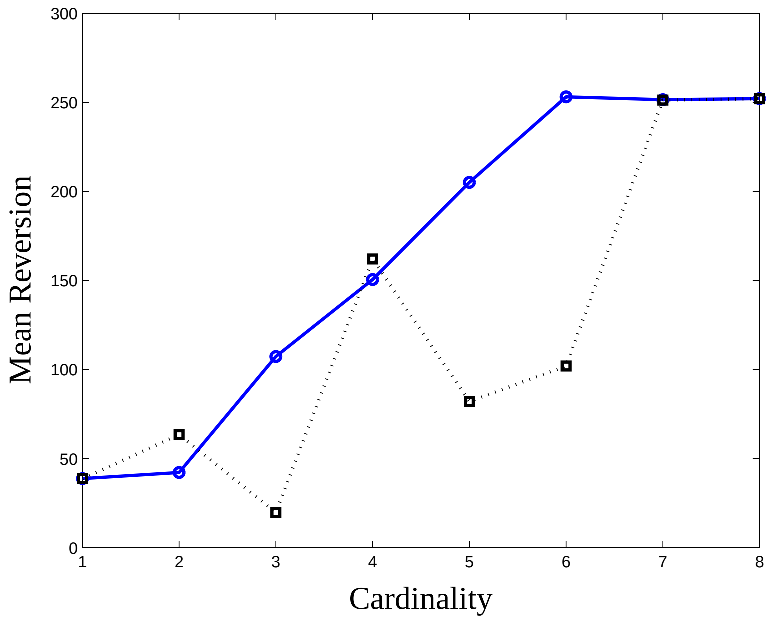

Figure 3: Mean reversion coefficient λ versus portfolio cardinality (number of nonzero coefficients) using the greedy search (solid line) and the semidefinite relaxation (dashed line) on U.S. swap rate data.

d'Aspremont, A., El Ghaoui, L., Jordan, M. & Lanckriet, G. R. G. (2007), 'A direct formulation for sparse PCA using semidefinite programming', *SIAM Review* **49**(3), 434–448.

Dickey, D. & Fuller, W. (1979), 'Distribution of the Estimators for Autoregressive Time Series With a Unit Root', *Journal of the American Statistical Association* **74**(366), 427–431.

Engle, R. & Granger, C. (1987), 'Cointegration and error correction: representation, estimation and testing', *Econometrica* **55**(2), 251–276.

Fama, E. & French, K. (1988), 'Permanent and Temporary Components of Stock Prices', *The Journal of Political Economy* **96**(2), 246–273.

Gatev, E., Goetzmann, W. & Rouwenhorst, K. (2006), 'Pairs Trading: Performance of a Relative-Value Arbitrage Rule', *Review of Financial Studies* **19**(3), 797.

Grossman, S. & Vila, J. (1992), 'Optimal Dynamic Trading with Leverage Constraints', *The Journal of Financial and Quantitative Analysis* **27**(2), 151–168.

Johansen, S. (1988), 'Statistical analysis of cointegration vectors', *Journal of Economic Dynamics and Control* **12**(2/3), 231–254.

Johansen, S. (1991), 'Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models', *Econometrica* **59**(6), 1551–1580.

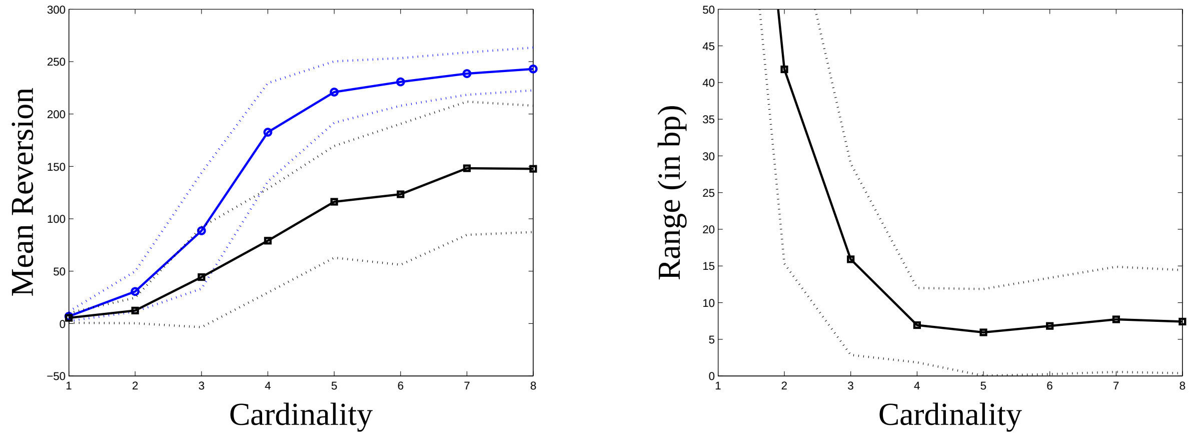

Figure 4: *Left:* mean reversion coefficient λ versus portfolio cardinality (number of nonzero coefficients), in sample (blue circles) and out of sample (black squares) on U.S. swaps. *Right:* out of sample portfolio price range (in basis points) versus cardinality (number of nonzero coefficients) on U.S. swap rate data. The dashed lines are at plus and minus one standard deviation.

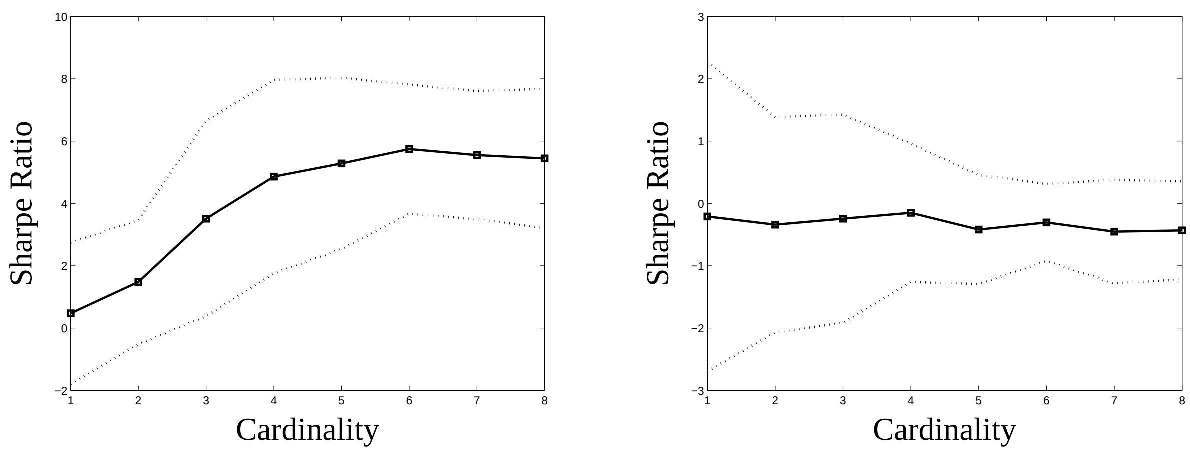

Figure 5: *Left:* average out of sample sharpe ratio versus portfolio cardinality on U.S. swaps. *Right:* idem, with a Bid-Ask spread of 1bp. The dashed lines are at plus and minus one standard deviation.

Jurek, J. & Yang, H. (2006), Dynamic portfolio selection in arbitrage, Technical report, Working Paper, Harvard Business School.

Kim, T. & Omberg, E. (1996), 'Dynamic Nonmyopic Portfolio Behavior', *The Review of Financial Studies* **9**(1), 141–161.

Liu, J. & Longstaff, F. (2004), 'Losing Money on Arbitrage: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities', *Review of Financial Studies* **17**(3).

Poterba, J. M. & Summers, L. H. (1988), 'Mean reversion in stock prices: Evidence and implications', *Journal of Financial Economics* **22**(1), 27–59.

Sturm, J. (1999), 'Using SEDUMI 1.0x, a MATLAB toolbox for optimization over symmetric cones', *Optimization Methods and Software* **11**, 625–653.

Toh, K. C., Todd, M. J. & Tutuncu, R. H. (1999), 'SDPT3 – a MATLAB software package for semidefinite programming', *Optimization Methods and Software* **11**, 545–581.

Wachter, J. (2002), 'Portfolio and Consumption Decisions under Mean-Reverting Returns: An Exact Solution for Complete Markets', *The Journal of Financial and Quantitative Analysis* **37**(1), 63–91.

Xiong, W. (2001), 'Convergence trading with wealth effects: an amplification mechanism in financial markets', *Journal of Financial Economics* **62**(2), 247–292.