Lecture Notes 10

Jussi Klemelä

December 3, 2014

1 The Maximization Problems

We consider single period portfolio selection. We have N basis assets with the single period returns $R^1_{t+1}, \ldots, R^N_{t+1}$. We define the optimality of the return distribution with the help of utility maximization. We consider set B_N of feasible portfolio vectors. Let

$$B_N \subset \left\{ (b^1, \ldots, b^N) : \sum_{i=1}^N b^i = 1 \right\}. \tag{1}$$

We choose the portfolio vector b^o_t which maximizes the expected utility:

$$b^o_t = \operatorname{argmax}_{b_t \in B_N} E_t u\,(b_t' R_{t+1}), \tag{2}$$

where $R_{t+1} = (R^1_{t+1}, \ldots, R^N_{t+1})'$.

Examples for the choice $B = B_N$ of the feasible portfolio vectors include the following.

1. Let us have two assets and take
$$B = \{(1,0),(0,1)\}.$$
Now we are investing everything into the first asset if $E_t u(R^1_{t+1})$ is larger than $E_t u(R^2_{t+1})$.

2. Let us have a stock whose return is R_{t+1} and the risk free rate whose return is r_t. Taking
$$B = \{(1,0),(-1,2)\}$$
means that we are either long of the stock, which gives return R_{t+1}, or we are short of the stock, which gives return $2r_t - R_{t+1}$.[1]

[1] Equivalently, we can consider two assets with returns R_{t+1} and $2r_t - R_{t+1}$. The first asset corresponds to being long of the stock and the second asset corresponding to being short of the stock. Then we take $B = \{(1,0),(0,1)\}$. This is also equivalent to considering returns $bR_{t+1} + (1-b)r_t$, where $b \in \{-1, 1\}$.

3. Let us have three assets and take

$$B = \{(0.5, 0, 5, 0), (0.5, 0, 0, 5), (0, 0.5, 0.5)\}.$$

This corresponds to the strategy where we put the equal weight to the two assets with the highest value for $E_t u(R_{t+1})$. We can generalize this to the strategy where we choose among N assets $M < N$ assets with the highest values for $E_t u(R_{t+1})$, and put equal weights to the M selected assets. Now,

$$B = \left\{ \left(\frac{1}{M} I_J(j) \right)_{j=1,\ldots,N} : J \subset \{1, \ldots, N\}, \#J = M \right\},$$

where $I_J(j) = 1$ if $j \in J$ and otherwise $I_J(j) = 0$, and we use the notation $(b_1, \ldots, b_N) = (b_j)_{j=1,\ldots,N}$.

We can also define the optimality of the return distribution using the variance penalized expected return. Now we define

$$b_t^o = \text{argmax}_{b_t \in B_N} \left[E_t \left(b_t' R_{t+1} \right) - \frac{\gamma}{2} \text{Var}_t \left(b_t' R_{t+1} \right) \right], \tag{3}$$

where $\gamma > 0$ is the risk aversion parameter.[2]

2 Prediction Approaches

When we want to find the portfolio vector maximizing the expected utility as in (2), then we have to estimate the expected utility transformed return $E_t u \left(b_t' R_{t+1} \right)$ for each portfolio vector $b_t \in B_N$. The notation E_t means that we take the expected value conditionally on the information available at time t. Estimation of the conditional expectation $E_t u \left(b_t' R_{t+1} \right)$ can be identified with the prediction of $u \left(b_t' R_{t+1} \right)$, when the best prediction is defined as the minimizer of the mean squared prediction error.

[2] We can also use excess returns to define the optimal portfolio vectors. This would lead to the replacement of (2) with

$$b_t^o = \text{argmax}_{b_t \in B_N} E_t u \left(b_t' R_{t+1} - r_t \right)$$

and to the replacement of (3) with

$$b_t^o = \text{argmax}_{b_t \in B_N} \left[E_t \left(b_t' R_{t+1} - r_t \right) - \frac{\gamma}{2} \text{Var}_t \left(b_t' R_{t+1} - r_t \right) \right],$$

where r_t is the risk free rate.

When we want to find the portfolio vector maximizing the variance penalized expected return as in (3), then we have to estimate the expected return $E_t\left(b_t' R_{t+1}\right)$ and the conditional variance $\mathrm{Var}_t\left(b_t' R_{t+1}\right)$ for each portfolio vector $b_t \in B_N$. The notation Var_t means that we take the variance conditionally on the information available at time t. Since $\mathrm{Var}_t(X) = E_t X^2 - (E_t X)^2$, the conditional variance is defined in terms of conditional expectations.

The prediction can be done using time space smoothing, state space smoothing, and other state space prediction methods.

2.1 Time Space Smoothing

In time space smoothing the next value Y_{t+1} of a time series $Y_1, \ldots, Y_t$ is predicted by a weighted average

$$\hat{f}(t+1) = \sum_{i=1}^{t} p_i Y_i, \tag{4}$$

where the weights satisfy $p_1 \leq \cdots \leq p_t$ with $\sum_{i=1}^{t} p_i = 1$. Here we denote with $\hat{f}(t+1)$ the estimator of the conditional expectation $E_t Y_{t+1}$. We can choose

$$p_i = \frac{\exp\{(i-t)/h\}}{\sum_{j=1}^{t} \exp\{(j-t)/h\}}, \tag{5}$$

where $h > 0$ is the smoothing parameter.

2.1.1 Expected Utility Optimizer

The general formula for the weighted average in (4) gives an estimator for $E_t u\left(b' R_{t+1}\right)$. The estimator is

$$\hat{f}_b(t+1) = \sum_{i=1}^{t} p_i u(b' R_i),$$

where we assume to have historical data $R_1, \ldots R_t$ of the returns of the portfolio components. The portfolio vector is then chosen as $\hat{b}_t = \mathrm{argmax}_{b \in B_N} \hat{f}_b(t+1)$.

2.1.2 Mean-Variance Optimizer

When we use mean-variance preferences to define the optimality of the return distribution, then we have to estimate

$$E_t\left(b' R_{t+1}\right) - \frac{\gamma}{2} \mathrm{Var}_t\left(b' R_{t+1}\right),$$

where $\gamma > 0$ is the risk aversion parameter. We take the estimate of $E_t\left(b'R_{t+1}\right)$ as

$$\hat{f}_b(t+1) = \sum_{i=1}^{t} p_i \cdot b'R_i$$

and the estimate of $\mathrm{Var}_t\left(b'R_{t+1}\right)$ as

$$\hat{g}_b(t+1) - \hat{f}_b(t+1)^2,$$

where

$$\hat{g}_b(t+1) = \sum_{i=1}^{T} p_i \left(b'R_i\right)^2.$$

2.2 State Space Prediction

The notation E_T means that the expectation is taken at time T, using information available at time T. If the available information is described by the vector X_T, then the expectation E_T can be taken as the conditional expectation

$$E_T u\left(b_T' R_{T+1}\right) = E\left[u\left(b_T' R_{T+1}\right) \mid X_T\right].$$

Define, for a fixed portfolio vector $b \in \mathbf{R}^N$ with $\sum_{i=1}^N b^i = 1$, the response variable

$$Y_{b,t} = u\left(b'R_{t+1}\right).$$

We assume that $(Y_{b,t}, X_t)$, $t = 1, \ldots, T-1$, are identically distributed, and denote by (Y_b, X) a random vector which has the same distribution as $(Y_{b,t}, X_t)$. Define the regression function

$$f_b(x) = E(Y_b \,|\, X = x), \qquad x \in \mathbf{R}^d.$$

Define the weight function $b : \mathbf{R}^d \to B$ by

$$b(x) = \mathrm{argmax}_{b \in B} f_b(x).$$

This function is estimated at time T by

$$\hat{b}_T(x) = \mathrm{argmax}_{b \in B} \hat{f}_{b,T}(x),$$

where

$$\hat{f}_{b,T} : \mathbf{R}^d \to \mathbf{R}$$

is a regression function estimate, constructed using data $(Y_{b,t}, X_t)$, $t = 1, \ldots, T-1$. At time T we choose the portfolio vector

$$\hat{b}_T(X_T).$$

We can use as explanatory variables the previous values of the returns of the basic assets. Then the vector of explanatory variables is

$$X_t = (R_t, \ldots, R_{t-k+1}),$$

where $k \geq 1$ is the lag parameter and $R_t = (R_t^1, \ldots, R_t^N)$ is the vector of the returns of the basix assets.

2.2.1 Linear Regression

In linear regression the regression function estimator is

$$\hat{f}_{b,T}(x) = \hat{\alpha} + \hat{\beta}' x,$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the least squares estimates, calculated with the regression data $(X_t, Y_{b,t})$, where $Y_{b,t} = u(b' R_{t+1})$, $t = 1, \ldots, T-1$. Now $\hat{f}_{b,T}(x)$ is the prediction for time $T+1$ utility transformed return, when $X_T = x$ is observed. We choose the portfolio weights $b_T = \text{argmax}_b \hat{f}_{b,T}(X_T)$.

2.2.2 Kernel Regression

In kernel regression the regression function estimate is

$$\hat{f}_{b,T}(x) = \sum_{t=1}^{T-1} p_t(x)\, Y_{b,t},$$

where

$$p_t(x) = \frac{K_h(X_t - x)}{\sum_{u=1}^{T-1} K_h(X_u - x)},$$

with $K_h(x) = K(x/h)$, $K : \mathbf{R}^d \to \mathbf{R}$ is the kernel function, and $h > 0$ is the smoothing parameter.