An LP formulation for portfolio optimization in Mavro

Nicos Christofides[1]
Alexander Christofides[1]
Simon Christofides[2]

1: Centre for Quantitative Finance,
Imperial College, London, SW7 2AZ, UK

2: Goldman Sachs

June 2007

1 Introduction

We start with assets $n \in N \equiv \{1, 2, \ldots, \hat{n}\}$. Consider equispaced time instances $t \in T^0 \equiv \{0, 1, 2, \ldots, \hat{t}\}$. Intervals $[t-1, t],\ t \in T \equiv T^0 - \{0\}$ of length Δ years are referred to as periods. The price of asset n at time $t \in T^0$ is $a_t(n)$ and its observed return in period $[t-1, t]$ is $r_t(n) \equiv \ln\left(\frac{a_t(n)}{a_{t-1}(n)}\right),\ t \in T$.

A portfolio $\mathbf{y} \equiv [y(1), y(2), \ldots, y(\hat{n})]^T$ is the allocation of $y(n)$ euros to asset $n \in N$. In general we assume a wealth of 1 euro, and take $y(n)$ as the fraction of wealth allocated to asset n. Given the above observations of asset returns, we want to find a portfolio $\mathbf{y}$ so that its performance over the 'future' period $[\hat{t}, \hat{t}{+}1]$ has some desirable characteristics. Let $\bar{r}(n)|_\Delta$ be the expected return of asset n over one period, and $c(n,m)|_\Delta$ be the covariance of the period-returns of assets n and m, both statistics derived from the above observations. Typically, these values are annualized by multiplying both with $\frac{1}{\Delta}$ to produce $\bar{r}(n)$ and $c(n,m)$ respectively. Assuming that the above statistical properties are constant over time (and are, therefore, applicable for the future period $[\hat{t}, t \hat{+} 1]$) Markowitz formulated the portfolio selection problem as:

$$\text{Maximize:} \quad z = \sum_{n \in N} \bar{r}(n).y(n) \tag{1}$$

subject to:

$$\sum_{n,m \in N} y(n)y(m)c(n,m) \leqslant \hat{\rho} \tag{2}$$

$$\sum_{n \in N} y(n) = 1 \tag{3}$$

$$y(n) \geqslant 0 \qquad \forall n \in N \tag{4}$$

Equation 1 is the expected portfolio return, the lhs of equation 2 is the variance of the portfolio return (which is the assumed measure of risk) and is required to be less than a given risk level $\hat{\rho}$, equation 3 is the budget constraint, and equation 4 excludes short positions. The matrix $[c(n,m)]$ is positive semi-definite, the optimization problem is convex and easily solvable as a quadratic program.

2 An LP formulation

An alternative measure of risk is the *mean absolute deviation* of the portfolio return introduced by Konno & Yamazaki :

$$\rho = \sum_{t \in T} q_t \left| \sum_{n \in N} (r_t(n) - \bar{r}(n))y(n) \right| \tag{5}$$

where $q_t = 1/\hat{t}$ is the probability that the asset return vector at time t is $[r_t(n), n \in N]$.

The modulus of an expression X, i.e. $|\mathrm{X}|$, can be rewritten as $(u+v)$ using two new variables $u \geqslant 0$ and $v \geqslant 0$, which satisfy:

$$u - v = \mathrm{X}, \qquad u.v = 0, \qquad u,\ v \geqslant 0$$

Thus, using ρ as the risk measure and the above trick to eliminate the modulus, the portfolio selection problem can be rewritten as

Problem P:

$$\text{Maximize:} \quad z = \sum_{n \in N} \bar{r}(n).y(n) \tag{6}$$

$$\text{subject to:} \quad \sum_{t \in T} q_t(u_t + v_t) \leqslant \hat{\rho} \tag{7}$$

$$u_t - v_t = \sum_{n \in N} (r_t(n) - \bar{r}(n))y(n), \qquad \forall t \in T \tag{8}$$

$$\sum_{n \in N} r_t(n).y(n) \geqslant -\check{\lambda} \qquad \forall t \in T \tag{9}$$

$$\sum_{n \in N} y(n) = 1 \tag{10}$$

$$u_t, \ v_t \geqslant 0, \qquad \forall t \in T \tag{11}$$

$$y(n) \geqslant 0, \qquad \forall n \in N \tag{12}$$

together with the constraints

$$u_t.v_t = 0, \qquad \forall t \in T$$

The theory of linear programming implies that these last constraints are redundant as pointed out by Chvatal. Thus the portfolio selection problem is defined by the pure LP given by equations 6 to 12.

In equation 7, $\hat{\rho}$ is now the upper limit of the mean absolute deviation risk measure, and equation 9 imposes that in the worst (observed) case t, the portfolio loss is not more than $\check{\lambda}$ (say 2% loss).

3 Enhancements to the LP in Mavro

3.1 Long only, or long/short portfolios

Substituting $y(n) = y(n)^+ - y(n)^-$ (where '+' implies a long and '-' a short position) in equations 6, 8, 9, and writing equations 10, 12, individually for $y(n)^+$ and $y(n)^-$, leads to an LP that allows equal long and short positions. The fact that $y(n)^+$ and $y(n)^-$ cannot both be positive for a given n (and which is implied by adding the constraints $y(n)^+.y(n)^- = 0, \forall n \in N$) need not be imposed explicitly.

3.2 Sector and/or country neutrality

It is a trivial matter to add several useful constraints to this model. For example, let S be the set of all sectors and N_s be the set of assets in sector

$s \in S$. Long/short sector neutrality can be imposed by the addition of constraints $\sum_{n \in N_s} [y(n)^+ - y(n)^-] = 0, \ \forall s \in S$. Similarly for country neutrality.

3.3 Portfolio turnover: Existing solution

If we already hold a portfolio $\mathbf{y}_{\text{old}}$, and want the new portfolio to be different from it only to an extent of, say, $100.\epsilon$ percent, we can impose this by adding the constraint:

$$\sum_{n \in N} |y(n) - y_{\text{old}}(n))| \leqslant \epsilon$$

We have already described the variable substitution that converts a modulus function to a linear expression and hence the above constraint does not complicate the LP model, but only slightly increases the number of variables.

3.4 Using forecasts

In the LP formulation, the time instances t need not be consecutive; indeed they do not have to be historical, or actual realizations., and don't even have to relate to 'time'. They can, for example, be scenarios generated as possibilities for the asset returns in the future time period $[\hat{t}, \hat{t}+1]$. Similarly, q_t does not have to be set to $1/\hat{t}$ for all observations and can be greater than this for the recent ones and less for the past ones. This means that forecasted values of the asset returns over period $[\hat{t}, \hat{t}+1]$ (say $\tilde{\mathbf{r}} = [\tilde{r}(1), \tilde{r}(2), \ldots, \tilde{r}(\hat{n})]$ can be added to (for example) the historical realizations above, with probability weight $\tilde{q}$ dependent on the confidence level given to this forecast. (Obviously, the rest of the probabilities then need to be adjusted.)

A shortcoming in the above simple addition of forecasted values, lies in the fact that the Mavro forecaster, forecasts asset returns individually with no account taken of their correlation. This means that the log-likelihood that the return vector $\tilde{\mathbf{r}}$ has been generated from the same stochastic process as the rest of the return vectors is small. However, we already have a procedure to *fully* correct this shortcoming as follows. Given that the observations $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{\hat{t}}$ are historical we can easily compute parameter vectors $\underline{\mu}$ and $\underline{\sigma}$ in an $\hat{n}$-factor GBM model.

The forecaster has nothing to say about $\underline{\sigma}$ and we keep this covariance matrix unchanged (computed using the shrinkage operator described in Ledoit & Wolf). However, by forecasting expected returns, the forecaster implies a different vector $\underline{\mu}$ over the next period $[\hat{t}, \hat{t}+1]$. (It is interesting to note here, that in parameter estimation, it is very well documented in the literature that by far the most unreliable estimate between $\underline{\mu}$ and $\underline{\sigma}$, is $\underline{\mu}$.) Let $\underline{\tilde{\mu}}$ be the forecasted value of $\underline{\mu}$.

In the note on the construction of a State Transition Graph, we described a general procedure which, given $\hat{n}$ assets following a joint GBM process with given parameter vectors/matrices $\underline{\mu}$ and $\underline{\sigma}$, we can generate $\hat{n}+1$ vectors $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{\hat{n}+1}$ with corresponding probabilities $p_1, p_2, \ldots, p_{\hat{n}+1}$, such that the mean values of the returns and their correlations fit exactly the values implied by the given $\underline{\mu}$ and $\underline{\sigma}$. Thus, although we cannot add the forecasted value of the asset returns directly to the historical returns, we can add the set of $\hat{n}+1$ vectors above, and achieve precisely what we want, namely a *local distortion* of the return distribution without a distortion of the asset interdependence.

3.5 Modifying the LP solutions

3.5.1 Portfolio size

The optimal solution $\mathbf{y}^*$ of the LP model is likely to contain money allocations (long or short) to many more assets than we wish to hold. Imposing a limit on this number requires the introduction of integer variables which leads to a MIP that is difficult to solve. Instead, we use a typical heuristic procedure as follows:
(i) Solve the LP
(ii) If the number of assets with non-zero money allocation is less than the maximum allowable, stop; else
(iii) Let n' be the asset n with the smallest value of $|y(n)|$. Set $y(n) = 0$ and repeat from (i)

3.5.2 Alternative solutions

In the LP formulation, the risk parameters $\hat{\rho}$ and $\check{\lambda}$ represent the risk-appetite of the investor. We know that the lower the value of these parameters the more risk averse is the investor. However, we may want to solve the LP several times for different values of these parameters to get a few 'optimal' portfolios (for different degrees of risk) from which a final selection is made.
Similarly, parameter $\tilde{q}$ determines the importance given to the forecasts. $\tilde{q} = 0$ implies that the forecasts are ignored and only historical (or simulated) performance is used, whereas $\tilde{q} = 1$ implies that only the forecasted returns are used.