IPM
INFORMED PORTFOLIO MANAGEMENT

INSIGHTS I Q4 2009

Tail Risk Optimisation

By

Alex Gioulekas, IPM
Boualem Djechiche, Mathematics Department, Royal Institute of Technology, Stockholm

Introduction

When combining different strategies in a multi-strategy product, our objective is to maximise expected return while minimising drawdowns. Drawdowns are severe when a number of strategies suffer extreme losses at the same time. Therefore the strategies must be weighted in a fashion that minimizes the coincidence of extreme negative events.

The conventional portfolio optimisation method is to minimise mean variance for a given expected return assuming normal joint distribution of returns. In real life application, this approach generates sub-optimal portfolios because it does not account for the fat tails of the actual distributions and because it measures correlation primarily around the means of the return series. Concurrence of extreme negative events is given a very low probability in this approach.

We therefore need a portfolio design tool that models more accurately the negative tails of the distributions and the coincidence probability of extreme events.

In **Section 1** we describe such a methodology. We model the return structure via the most general form of the multivariate Student's t-distribution. Co-dependence is described by the t-copula. It turns out that this solution is computationally intensive and thus impractical.

To remedy this situation, we have developed a "Practical Solution". We use the simplest form of the multivariate t-distribution to describe how returns are jointly distributed and rescale the data so that fat negative tails are penalised. This is described in **Section 2**.

To demonstrate that the practical solution indeed yields similar results to the exact solution, we devise a thought experiment in **Section 3.**

IPM has been applying this methodology since 2005 to optimise its GTAA portfolios. We show how these portfolios have performed in real life compared to hypothetical portfolios where the GTAA strategies were equally weighted in **Section 4.**

In **Section 5** we show the results of applying the methodology to long-only equity index portfolios. We conclude that tail-risk optimisation can protect a portfolio of diversified strategies by avoiding concurrence of extreme losses in a subset of the strategies but cannot perform miracles in a portfolio where all strategies use similar alpha sources and underperform at the same time.

All relevant mathematical formulas of the analysis can be found in the Appendix.

1. Minimise drawdown: an exact solution

In portfolio management the perfect storm scenario is a turn of events when a large subset of strategies or assets suffers large losses simultaneously.

When designing portfolios it is important to have a methodology that indentifies which strategies suffer large losses at the same time and thus allocates smaller weights to those strategies.
We have developed such a methodology. Our task is to find the optimal portfolio weights able to minimise the multi-period portfolio drawdown.

The first step in this methodology is to describe the return structure of the strategies. To this end we use the most general form of the multivariate Student's t-distribution (Appendix A). In one dimension, the t-distribution has fatter tails than the normal distribution: extreme events are more likely to happen. The fatness of the tails in the t-distribution is described by the parameter "degrees of freedom" (dof). The lower the dof, the fatter the tails. As the dof goes to infinity, the t-distribution becomes the normal distribution[1].

The dependence between the marginal distributions[2] is described by the t-copula. What is a copula? It is a function for creating a joint probability distribution for two or more marginal distributions (Appendix A). For details see Ref[2] and Ref[3]. The advantage of using a multivariate t-distribution is that tail events (and we are mainly concerned about negative tail events) occur and coincide more frequently than in the multivariate normal distribution. We fit the data in two steps:

- First we consider each strategy and fit a marginal distribution to the strategy's return series. We concentrate on the negative returns and calculate the dof that best matches the negative returns of each strategy (Appendix B).

- Then we describe the return dependence. We calculate the dof of the t-copula that best fits the joint dispersion of the data – Appendix C describes how this is done using a maximum likelihood device. We use the rank correlation matrix to calculate the covariance matrix in the t-copula. This is a more robust method for calculating correlation because it weighs all events equally and does not only emphasize dependence around the mean.

We are now ready to optimise the weights of the multi-strategy portfolio. As there is no closed form solution to this problem, we need to perform the following computations:

- Simulate a large number of outcomes drawn from our multi t-distribution, whose copula has the estimated dof and each marginal is t-distributed with its own dof (Appendix D).

- Based on these, simulations compute the optimal weights that minimise the multi-period drawdown. We wrote a linear programming algorithm to solve this optimisation problem.

But then we run into trouble: the task of optimising a portfolio with more than four strategies exceeds computer capacity.

---

[1] An equity index in the developed world, typically has dof=4 or higher. Bonds and most currencies have thinner tails and thus higher dof.

[2] The term marginal distribution refers to the return distribution of each strategy considered separately.

2. Minimise tail risk: a practical solution

To overcome this trouble we have developed an approximate solution, which is fast to compute and should approximate the results of the exact solution.

What led us to this discovery?
We knew the analytic solution to a simpler problem: given t-distributed strategies with the *same* degree of freedom and a positive-definite dispersion matrix, find the minimum Value-at-Risk portfolio or equivalently the portfolio with the minimum *one-period* drawdown. We also noted that the Value-at-Risk is proportional to the square root of the determinant of the dispersion matrix.

The first step is to use a simplified multi t-distribution, where the copula and every marginal have the same dof, and to find the dof of the copula that best matches the correlation structure of the data[3]. This approximation is crude: all marginals have the same dof and therefore similar kurtosis.

The next step is to account for the idiosyncrasy of each return series. Given that minimising Value-at-Risk is equivalent to minimising dispersion, we make strategies with fat tails appear to have higher volatility. We transform each return series as follows:

- First, we take the mean out.

Then, we divide by the square root of the dof, as determined by fitting a univariate t-distribution to the strategy's returns. Fat tails (small dof) are transformed to high standard deviation. To see why we divide the returns by the square root and not any other function of the dof, take a look at the formula (3) for the density of the multivariate t-distribution: the dof times the dispersion is dimensionless. This means that standard deviation and the inverse of the square root of the dof play similar roles in determining the shape of the above distribution.

- Finally, we add back the mean to the rescaled series. The strategies' expected return is thus maintained.

---

[3] To estimate the dof we maximize the logarithm of the distribution's likelihood function with respect to dof.

3. A thought-experiment

In order to demonstrate that the practical solution yields the same results as the exact solution, we performed the following thought-experiment in the beginning of 2005, when we first developed the methodology.

We first optimise a portfolio of US, UK, Japan, and Germany MSCI total return indices, by using a mean variance optimisation. We then replace the US index returns by a fictional series designed to:

- Have the same mean return and volatility as the US index and the same linear correlation with the remaining assets.

- Have a succession of large losses towards the end of the period resembling a bursting bubble, see Figure 1. The dof for this fictional asset is 2.5, compared to 4.5 for Japan and 4 for US and Germany and has a large negative skew.

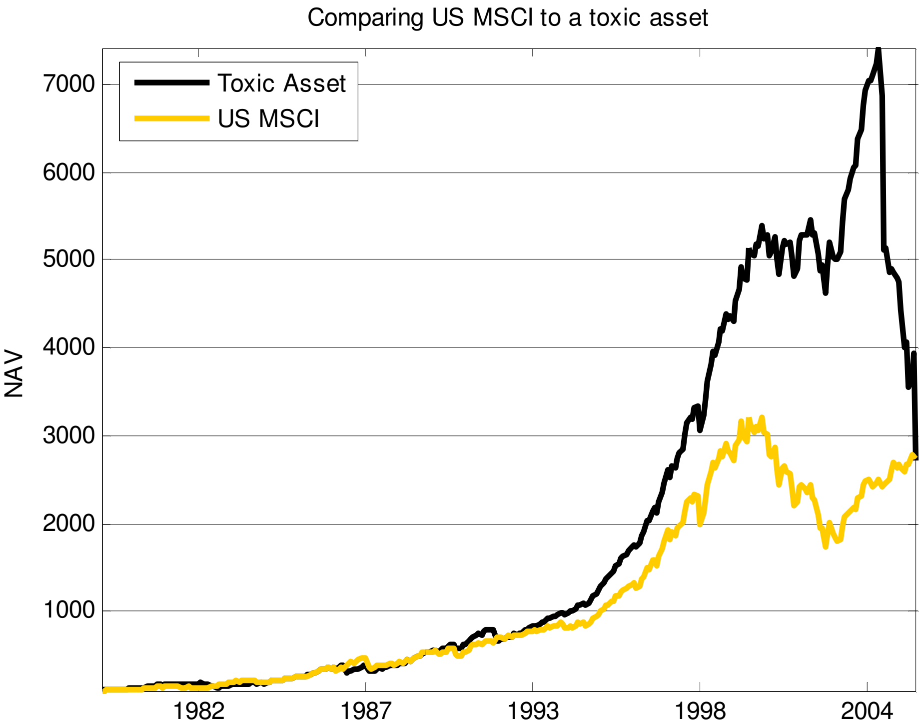

Figure 1. The fictitious asset: a bubble that bursts.

Then we calculate the weights of four portfolios comprising the fictional index and the UK, Japan and Germany MSCI indices:

- The minimum variance portfolio
- The optimal portfolio derived from the practical method of section 2 – which we shall call the tail-risk optimal portfolio
- The minimum shortfall portfolio derived from the numerical method of section 1 and
- The minimum multi-period drawdown portfolio derived from the numerical method of section 1

The results are shown in Figure 2 below.

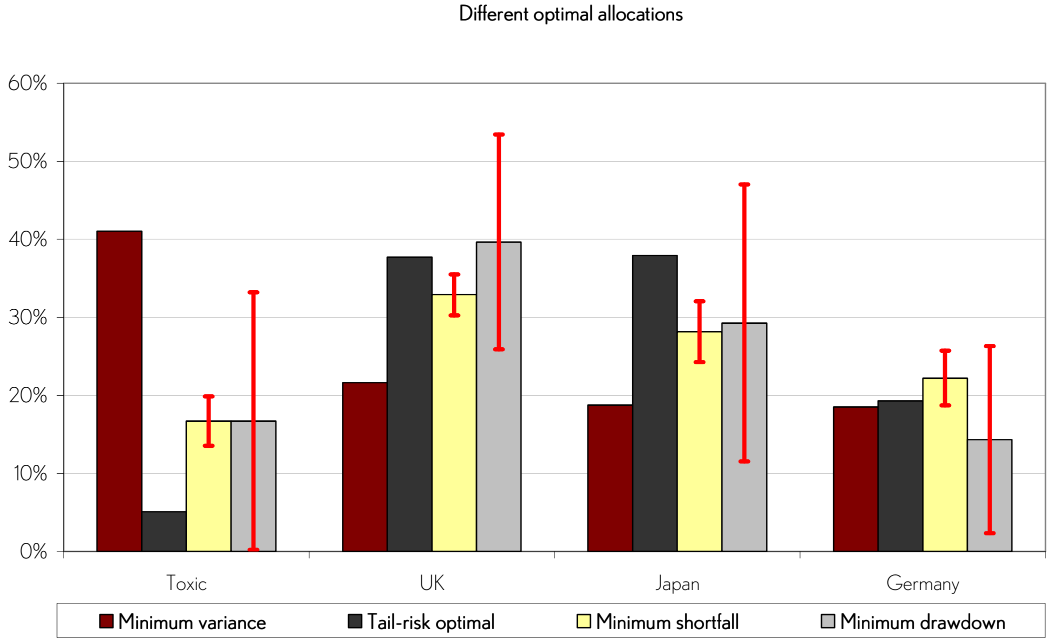

Figure 2. Comparison of portfolios obtained by different optimisation methodologies. The weights are calculated using monthly returns from 1980 to 2005.

The mean variance minimisation gives the fictional asset the same weight it gave the US MSCI index. This optimisation method is oblivious to the succession of extreme negative returns in the toxic asset, because it focuses on mean returns and correlation around the means of the distribution and misses the negative tail events.

The practical solution errs on the side of caution and gives the toxic asset a smaller weight compared to the portfolio with the minimum shortfall and to the portfolio with the minimum multi-period drawdown.

The last two portfolios are obtained numerically as described in section 1. We generate 10 scenarios by drawing 2000 observations per scenario from the multivariate t-distribution fitted to the 4 asset returns. For every scenario we find the optimal weights and then we calculate the average over the 10 scenarios. The error bars around the average weight indicate how the weight varies from simulation to simulation and correspond to the 95% confidence level. There is greater uncertainty surrounding the weights in the optimisation that minimises multi-period drawdowns.

It took 3 hours to compute the exact solution using the best technology available in 2005 to solve this type of linear programming problems. We estimated that it would have taken 1 week to find the optimal portfolio for a scenario with 10 000 observations.

This thought experiment is not a proof that the practical solution always gives similar results to the exact solution, but demonstrates that the practical solution does what was designed to do: identify and penalize fat tails.

4. Tail optimisation in real life

Our GTAA model is based on four strategies: currency, country bond, country stock and asset class selection. These strategies have similar information ratios and are designed to have linear correlation close to zero. In the beginning of 2005 we used the tail optimisation method to weigh the four strategies, see Ref[1], available upon request. The weights determine how the risk budget is strategically allocated to the four processes. Because the last two strategies exhibit considerable tail correlation (tail correlation coefficient=0.32, see Ref[1]), the tail optimisation assigned to the strategies weights 1/3, 1/3, 1/6 and 1/6 respectively, instead of allocating a quarter of the risk budget to each.

Since 2005 we have been updating the results of the optimization every year by using the 12 extra monthly observations to expand the return series and better define their tail dependence. These updates have taken some risk away from the currency and the asset class selection and allocated it to the stock and bond relative value processes.

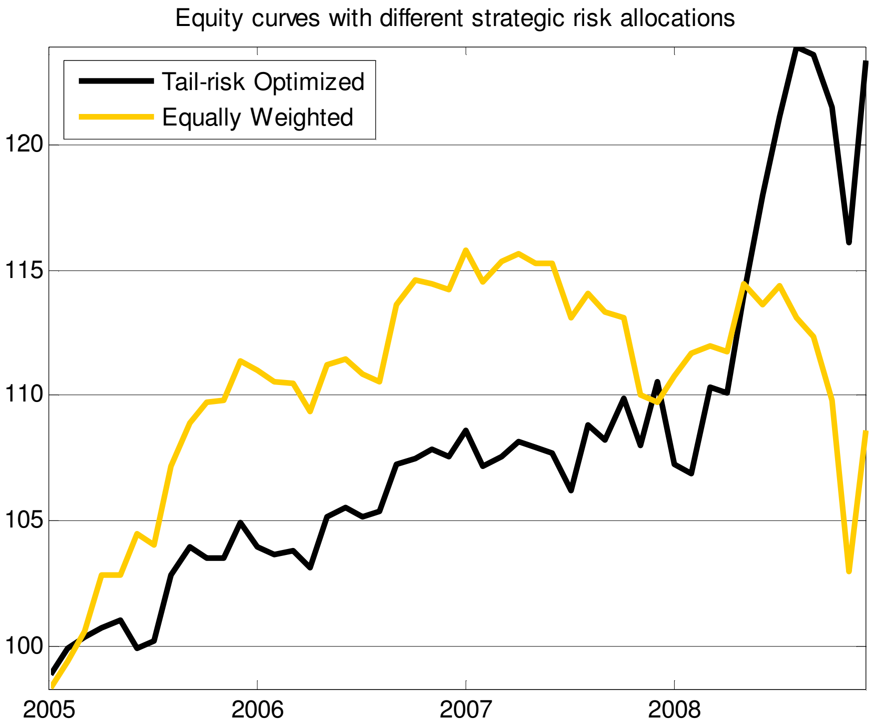

Figure 5. Real life performance of a tail-risk optimised portfolio compared to performance of a hypothetical portfolio with equally divided strategic risk.

Figure 5 compares the performance of the actual setup to a hypothetical setup that divides the strategic risk equally to the four strategies. Note that these return series are based on raw model output and do not include our risk management overlay. Our actual returns, based on the tail-risk optimized output, are superior to the latter.

Figure 6 compares the drawdowns between the actual and the equally weighted portfolio.

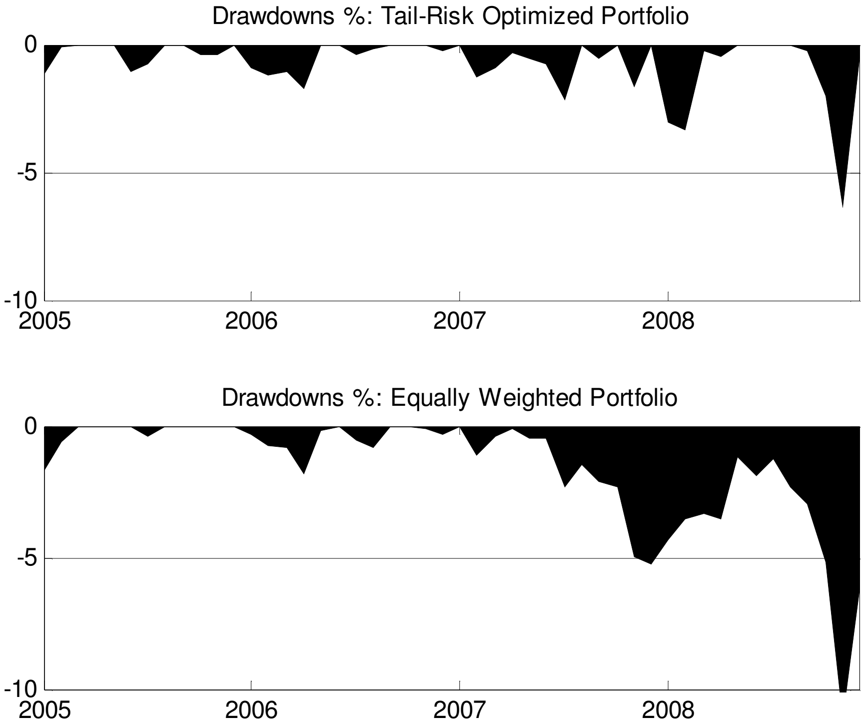

Figure 6. Real life drawdowns of a tail-risk optimised portfolio compared to drawdowns of a hypothetical portfolio with equally divided strategic risk

5. Tail optimisation of a stock portfolio

We are confident that tail optimisation unveils and penalises tail dependence between strategies. But how potent is it in preventing large drawdowns when all portfolio components are assets or strategies that fall and rise at the same time?

This situation arises in the example we consider next. We wish to optimise a long-only portfolio comprising the total return MSCI indices for the US, UK, Japan, Germany and France from the standpoint of a US investor who hedges the currency exposure.

We compare three methods: the tail-risk optimisation, equally weighting and variance minimisation.
Starting on the last day of 1998, we calculate the portfolio weights every three years. We base the calculation on monthly data spanning the period January 1979 to March 2009.

Figure 7. Weights of the tail-risk optimal portfolio calculated on different occasions.

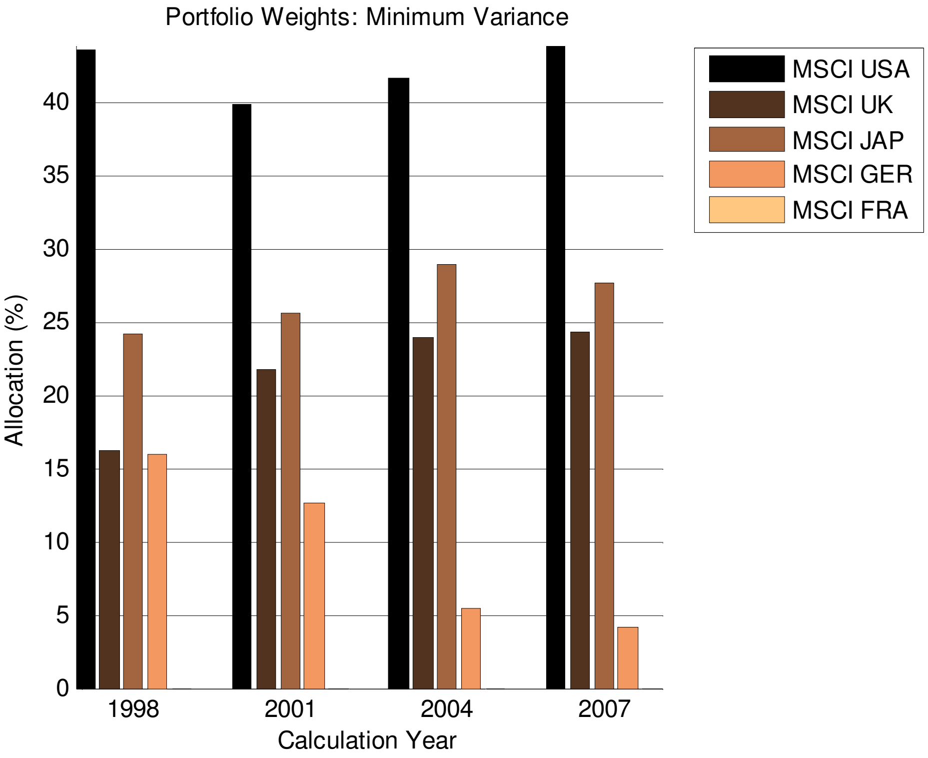

Figure 8. Weights of the tail-risk optimal portfolio calculated on different occasions.

Two things stand out when comparing the tail optimised to the minimum variance portfolios: The portfolio composition is different but performance is similar.

Figure 7 shows that the tail optimisation avoids the UK and Germany mainly because those return distributions are fat-tailed.

Figure 8 shows that minimum variance excludes France. France has a higher information ratio and similar volatility compared to Germany, but it has higher correlation to the other assets than Germany.

Not surprisingly the performance of the three portfolios is similar. Some gains up to 2007 have been erased in the recent bear market.

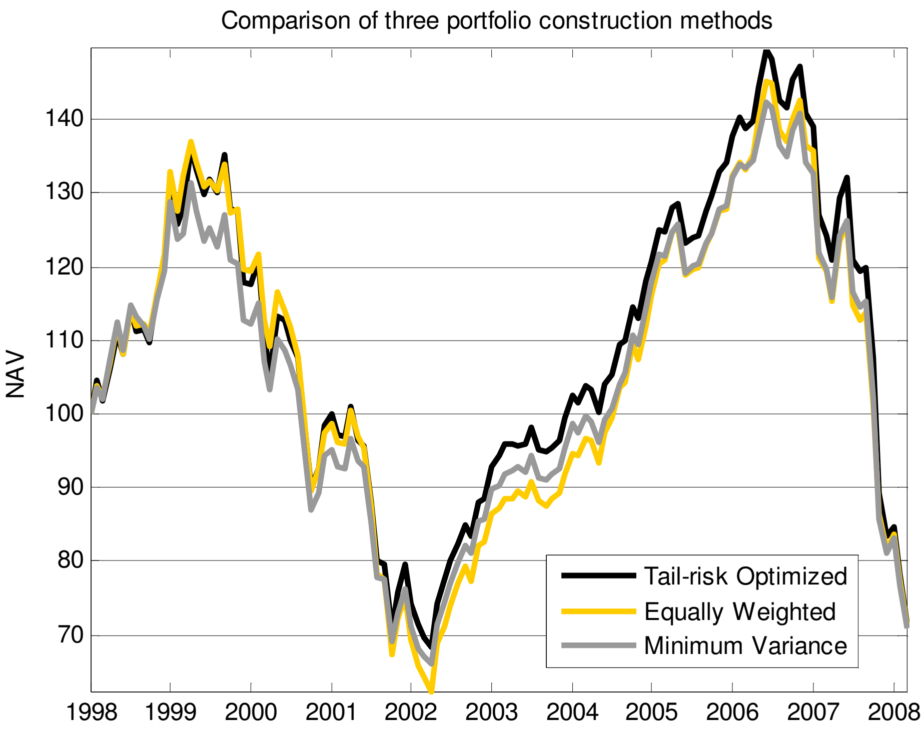

Figure 9. Three equity index portfolios constructed using different methodologies show remarkably similar performance.

Conclusion

We have developed a practical method for minimising multi-period draw-downs in a multi-strategy portfolio by better modelling the probability of "perfect storm" scenarios when a number of strategies exhibit simultaneously large losses.

For four strategies we are able to give a more accurate solution to the problem based on scenario generation and optimisation via a linear programming algorithm. For more than four strategies or assets this method is impractical with today's technology.

The application of the tail risk optimisation in real life gives our portfolios an edge vis-a-vis simpler design approaches.

The example of the long-only stock index portfolio demonstrates that portfolio robustness depends primarily on the quality and diversity of the underlying strategies. Only then the optimisation tools can produce a robust portfolio by avoiding coincidence of loss in a subset of strategies. If the alpha sources underlying the strategies are related, no optimisation method can prevent the "perfect storm".

Appendix & References

A: Some facts about the multivariate t-distribution

The one-dimensional t-distribution, with ν degrees of freedom and zero mean has probability density

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu})} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \qquad (1)$$

The variance of the t distribution is defined only if the degrees of freedom is larger than 2

$$\mathrm{var}(X) = \frac{\nu}{\nu - 2}$$

The most general form of the multivariate t-distribution has probability density

$$F(x_1, \ldots, x_n) = C_{R,\nu}(t_{\nu_1}(x_1), \ldots, t_{\nu_n}(x_n)) \qquad (2)$$

Note that the t-copula $C_{R,\nu}$ and the marginals $t_{\nu_i}(x_i)$ have different degrees of freedom.

The simplest form of the multivariate t- distribution, used in the model of section 3, is one where the copula and each marginal have the same degrees of freedom.

The n-dimensional vector $\mathbf{X} = (X_1, \ldots, X_n)$ has a multivariate t-distribution with ν degrees of freedom, mean vector μ and a positive-definite dispersion matrix Ω, if its probability density is

$$f(\mathbf{x}) = \frac{\Gamma(\frac{\nu+n}{2})}{\Gamma(\frac{\nu}{2})\sqrt{(\pi\nu)^n \det{(\Omega)}}} \left(1 + \frac{(\mathbf{x} - \mu)^t \Omega^{-1} (\mathbf{x} - \mu)}{\nu}\right)^{-\frac{\nu+n}{2}} \qquad (3)$$

B: Fitting the marginal distributions to the data

To fit the actual returns to known distributions we use the so-called quantile-quantile plot: we compare the quantiles of the data to the quantiles of different known distributions by plotting the former against the latter. For our purposes, we fit the actual return distributions to the centered symmetric student $t(\nu)$distributions with ν degrees of freedom.

The results of the q-q plots for two of the strategies underlying our GTAA model are given in the figures below.

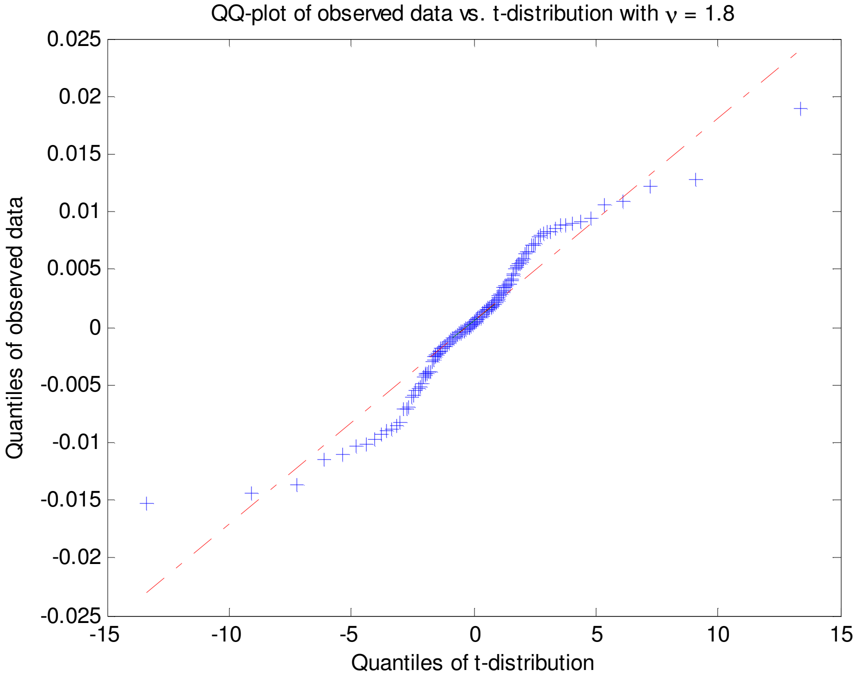

Figure B-1. The lower tail of the asset class selection returns fitted with the t(1.8) distribution.

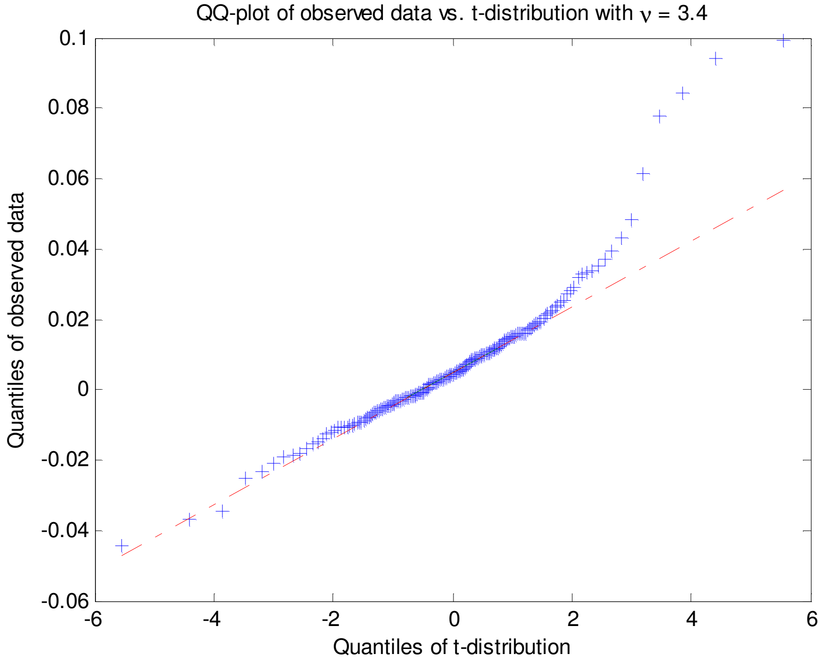

Figure B-2. The lower tail of the currency selection returns fitted with the t(3.4) distribution.

C: Fitting the copula to the data dispersion structure

To estimate the degrees of freedom ν of the multivariate t-distribution from the data correlation matrix R, we use the maximum likelihood methodology. This estimate is obtained by maximizing the "log-likelihood" function

$$\log L(\nu, \Omega; (\hat{X}_{\cdot,1}, \hat{X}_{\cdot,2}, \hat{X}_{\cdot,3}, \hat{X}_{\cdot,4}) = \sum_{i=1}^{N} \log f((\hat{X}_{i,1}, \hat{X}_{i,2}\hat{X}_{i,3}, \hat{X}_{i,4}))$$

where f is the density of the multivariate t distribution and $\hat{X}_{\cdot,i}$ are the column vectors of outcomes of the marginals X_i.

The density of a t-copula associated with a correlation matrix R has the form

$$c^t_{R,\nu}(u_1, \ldots, u_n) = \frac{1}{\sqrt{\det(R)}} \frac{\Gamma(\frac{\nu+n}{2})\Gamma(\frac{\nu}{2})^{n-1}}{\Gamma(\frac{\nu+1}{2})^n} \frac{\prod_{k=1}^n \left(1 + \frac{y_k^2}{\nu}\right)^{\frac{\nu+1}{2}}}{\left(1 + \frac{\mathbf{y}^t R^{-1} \mathbf{y}}{\nu}\right)^{\frac{\nu+n}{2}}}$$

where $\mathbf{y} = (y_1, \ldots, y_n)$ with $y_k = t_\nu^{-1}(u_k), \; u_k \in [0,1]$.

D: Simulating from the t-distribution

Formula (3) provides a straightforward algorithm for generating random scenarios from the general t-distribution (called meta t-distribution). Given a t-copula $C_{R,\nu}$ and t-distributed marginals with different degrees of freedom $\nu_i, \ i = 1, \dots, n$.

- Find the Cholesky decomposition A of R.
- Simulate n independent random variables $Z_1, \dots, Z_n$ from $N(0,1)$.
- Simulate a random variate S from χ^2_ν independent of $Z_1, \dots, Z_n$.
- Set $\mathbf{Y} = \frac{\sqrt{\nu}}{\sqrt{S}} A\mathbf{Z}$, with $\mathbf{Z} = (Z_1, \dots, Z_n)$.
- Set $U_i = t_{\nu_i}(Y_i)$ for $i = 1, \dots, n$.
- The vector $(t^{-1}_{\nu_1}(U_1), \dots, t^{-1}_{\nu_n}(U_n))$ has the same probability distribution as $\mathbf{X}$.

E: Minimizing the one-period drawdown

In Section 3 we solve the following optimization problem:
Given a loss level $L > 0$ and a confidence level α, find the portfolio weights u that maximize the expected return of the portfolio, while limiting below α the probability that the drawdown exceed the level $-L$

$$\max\{\mu_p \ \text{ such that } \ \mathbf{P}\left(X_p \le -L\right) \le \alpha, \ \mathbf{1}^t u = 1\} \qquad (4)$$

If the returns are t-distributed with ν degrees of freedom $\mathbf{X} \sim t_n(\nu, \mu, \Omega)$, then each component $X_i \sim t_1(\nu, \mu_i, \omega_i^2)$, where $\omega_i^2 = \Omega_{ii}$ and the total portfolio is also t distributed

$$X_p = u^t \mathbf{X} \sim t_1(\nu, u^t\mu, u^t\Omega u).$$

Rearranging, we get that

$$T = \frac{X_p - \mu_p}{\omega_p} \sim t_1(\nu, 0, 1)$$

is the standard centered symmetric t- distribution, where $\mu_p = u^t\mu$ and $\omega_p^2 = u^t\Omega u$.

Thus the problem (4) is equivalent to

$$\max\{\mu_p \ \text{ such that } \ \mathbf{P}\left(T \le k_a\right) \le \alpha, \ \mathbf{1}^t u = 1\} \qquad (5)$$

The expression $\mathbf{P}(T \le k_\alpha) = \alpha$ means that the k_α - quantile of T is equal to α.

Obviously what quantile of the t-distribution is equal to α depends only on the degrees of freedom of the multivariate t-distribution and is given by the inverse of the t cumulative distribution function (e.g. MATLAB function tinv).

The probability of drawdown $\mathbf{P}\left(X_p \le -L\right) = \alpha$ is in fact the $(1-\alpha)$-level Value-at-Risk for our t-distribution and thus the optimization problem (4) is equivalent to minimizing the Value-at-Risk at the $(1-\alpha)$-level.
This is why. Let $L = VaR_\alpha$, then

$$\mathbf{P}\left(X_p \le -VaR_\alpha\right) = \alpha \quad \text{if and only if} \quad VaR_\alpha = -\mu_p - k_\alpha\omega_p$$

Therefore problem (4) is equivalent to the following one:

$$\max\{\mu_p \text{ such that } VaR_\alpha = -\mu_p - k_\alpha \omega_p,\ \mathbf{1}^t u = 1\}$$

or equivalently

$$\min\{-\mu_p - VaR_\alpha = -VaR_{1-a},\ \mathbf{1}^t u = 1\}.$$

The analytic solution to this problem and its dual (4), is given in Ref[1].

When we optimize a portfolio of strategies we allow for leverage and we do not have the budget constraint. The optimization problem is:

Given a return constraint μ_{exp} and a confidence level α, find the portfolio weights that minimize the Value-at-Risk of the portfolio at the $(1-\alpha)$ level

$$\min\{VaR = -u^t\mu - k_a\sqrt{u^t\Omega u}\} \text{ such that } u^t\mu \geq \mu_{exp},\ u_{min} \leq u \leq u_{max}$$

where as before k_a is the α-quantile of the multivariate t-distribution fitted to the data.

References

[1] Djehiche, B. & Gioulekas, A. (2004). How to combine strategies while keeping drawdowns at a minimum. *IPM Internal Report.*

[2] Embrechts, P., Frey, R. & McNeil, A. (2005). Quantitative Methods Risk Management. *Princeton Series in Finance.* Princeton.

[3] Fan, H., Fang, K. & Kotz, S. (2002). The meta-elliptic distributions with given marginals. *J. Multivariate Anal. 82,* 1-16.