【仕事関連】Asset Allocation ModelをRで組んでみた。

AyatoAshihara

2019-04-25

今回はいつもと違って金融関連の記事を書きました。というのも社内でワークショップがあり、アセットアロケーションモデルを構築する必要が生じたからです。休日は国会図書館に籠り、先行研究を漁った結果、分散共分散行列をGARCHやARMAで推定し、最小分散ポートフォリオの性能向上を目指している論文がいくつか存在することがわかりました。今回はこれをRで実践し、推定方法によってパフォーマンスがどれほど変化するのかを実証してみたいと思います。

おはこんばんにちは。勤め先で、アセットアロケーションに関するワークショップに参加したので、この分野は完全なる専門外ですがシミュレーションをしてみたいと思います。今回は、最小分散ポートフォリオ(minimum variance portfolio)を基本ポートフォリオとしたうえで、その分散共分散行列（予測値）をどのように推計するのかという点について先行研究を参考にエクササイズしていきたいと思います。先行研究は以下の論文です（オペレーションリサーチのジャーナルでした）。

[<https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2947643:embed:cite>]

### 最小分散ポートフォリオ

最小分散ポートフォリオの詳しい説明はここでは割愛しますが、要は各資産（内株、外株、内債、外債、オルタナ）のリターンの平均と分散を計算し、それらを縦軸平均値、横軸分散の二次平面にプロットしたうえで、投資可能範囲を計算し、その集合の中で最も分散が小さくなるポートフォリオの事らしいです（下図参照）。

[f:id:osashimix:20190217131412p:plain]

先行研究のCarroll et. al. (2017)では、「 this paper focusses on minimum-variance portfolios requiring only estimates of asset covariance, hence bypassing the well-known problem of estimation error in forecasting expected asset returns. 」と記載されており、現状でも期待リターンの推計は難しく、それを必要としない最小分散ポートフォリオは有益で実践的な手法であるといえます。最小分散ポートフォリオの目的関数は、その名の通り「分散を最小化すること」です。今、各資産のリターンを集めたベクトルを[tex:r]、各資産の保有ウェイトを、ポートフォリオリターンをで表すことにすると、ポートフォリオ全体の分散は以下のように記述できます。

ここではの分散共分散行列です。よって、最小化問題は以下になります。

$$
\min\_{\theta}(\theta^{T}\Sigma\theta) \\
s.t 1^{T}\theta = 1
$$

ここでは、フルインベストメントを制約条件に加えています。ラグランジュ未定乗数法を用いてこの問題を解いてみましょう。ラグランジュ関数は以下のようになります。

1階の条件は、

$$
\displaystyle\frac{\partial L}{\partial \theta} = 2\Sigma\theta + 1\lambda = 0 \\
\displaystyle \frac{\partial L}{\partial \lambda} = 1^{T} \theta = 1
$$

1本目の式をについて解くと、

となります。ここで、です。これを2本目の式に代入し、について解きます。

$$
1^{T}\Sigma1\lambda^{\*} = 1 \\
\displaystyle \lambda^{\*} = \frac{1}{1^{T}\Sigma^{-1}1}
$$

だったので、を消去すると、

となり、最適なウェイトを求めることができました。とりあえず、これをRで実装しておきます。

# Global Mimimum Variance Portfolio  
gmv <- function(r\_dat,rcov){  
 library(MASS)  
 i <- matrix(1,NCOL(r\_dat),1)  
 r\_weight <- (ginv(r\_cov)%\*%i)/as.numeric(t(i)%\*%ginv(r\_cov)%\*%i)  
 wr\_dat <- r\_dat\*as.numeric(r\_weight)  
 portfolio <- apply(wr\_dat,1,sum)  
 pr\_dat <- data.frame(wr\_dat,portfolio)  
 sd <- sd(portfolio)  
 result <- list(r\_weight,pr\_dat,sd)  
 names(result) <- c("weight","return","portfolio risk")   
 return(result)  
}

入力は各資産のリターンと分散共分散行列になっています。出力はウェイト、リターン、リスクです。

### 分散共分散行列をどのように求めるか

最小分散ポートフォリオの計算式は求めることができました。次は、その入力である分散共分散行列をどうやって求めるのかについて分析したいと思います。一番原始的な方法はその時点以前に利用可能なリターンデータを標本として分散共分散行列を求め、その値を固定して最小分散ポートフォリオを求めるというヒストリカルなアプローチかと思います（つまりウェイトも固定）。ただ、これはあくまで過去の平均値を将来の予想値に使用するため、いろいろ問題が出てくるかと思います。専門外の私が思いつくものとしては、前日ある資産Aのリターンが大きく下落したという場面で明日もこの資産の分散は大きくなることが予想されるにも関わらず、平均値を使用するため昨日の効果が薄められてしまうことでしょうか。それに、ウェイトを最初から変更しないというのも時間がたつにつれ、最適点から離れていく気がします。ただ、ではどう推計するのかついてはこの分野でも試行錯誤が行われているようです。Carroll et. al. (2017)でも「The estimation of covariance matrices for portfolios with a large number of assets still remains a fundamental challenge in portfolio optimization.」と述べられていました。この論文では以下のようなモデルを用いて推計が行われています。いずれも、分散共分散行列を時変としているところに特徴があります。

1. Constant conditional correlation (CCC) model

元論文はこちら。

[<https://www.jstor.org/stable/2109358?read-now=1&seq=3#page_scan_tab_contents:title>]

まず、分散共分散行列と相関行列の関係性から、となります。ここで、は相関行列、はで各資産期の標準偏差を対角成分に並べた行列です。ここから、とを分けて推計していきます。まず、ですが、こちらは以下のような多変量GARCHモデル(1,1)で推計します。

$$
r\_{t} = \mu + u\_{t} \\
u\_{t} = \sigma\_{t}\epsilon \\
\sigma\_{t}^{2} = \alpha\_{0} + \alpha\_{1}u\_{t-1}^{2} + \alpha\_{2}\sigma\_{t-1}^{2} \\
\epsilon\_{t} = NID(0,1) \\
E(u\_{t}|u\_{t-1}) = 0
$$

ここで、はリターンの標本平均です。は推定すべきパラメータ。をGARCHで推計しているので、リターンの分布が正規分布より裾野の厚い分布に従い、またリターンの変化は一定ではなく前日の分散に依存する関係をモデル化しているといえるのではないでしょうか。とりあえずこれでの推計はできたということにします。次にの推計ですが、このモデルではリターンを標本として求めるヒストリカルなアプローチを取ります。つまり、は定数です。よって、リターン変動の大きさは時間によって変化するが、各資産の相対的な関係性は不変であるという仮定を置いていることになります。

1. Dynamic Conditional Correlation (DCC) model 元論文はこちら。

[<http://www.cass.city.ac.uk/__data/assets/pdf_file/0003/78960/Week7Engle_2002.pdf>]

こちらのモデルでは、を求めるところまでは①と同じですが、[の求め方が異なっており、ARMA(1,1)を用いて推計します。相関行列はやはり定数ではないということで、期までに利用可能なリターンを用いて推計をかけようということになっています。このモデルの相関行列は、

です。ここで、は期での条件付分散共分散行列で以下のように定式化されます。

ここで、はヒストリカルな方法で計算した分散共分散行列であり、はパラメータです。この方法では、先ほどとは異なり、リターン変動の大きさが時間によって変化するだけでなく、各資産の相対的な関係性も通時的に変化していくという仮定を置いていることになります。金融危機時には全資産のリターンが下落し、各資産の相関が正になる事象も観測されていることから、この定式化は魅力的であるということができるのではないでしょうか。

1. Dynamic Equicorrelation (DECO) model 元論文はこちら

[<https://faculty.chicagobooth.edu/bryan.kelly/research/pdf/deco.pdf>]

この論文はまだきっちり読めていないのですが、相関行列[tex:R\_{t}]の定義から

となるようです。ここで、はスカラーでequicorrelationの程度を表す係数です。equicorrelationとは平均的なペアワイズ相関の事であると理解しています。つまりは欠損値がなければ普通の相関と変わりないんじゃないかと。ただ、資産が増えればそのような問題にも対処する必要があるのでその点ではよい推定量のようです。は以下のように求めることができます。

ここで、はN×1ベクトルで要素は全て1です。また、はのi,j要素です。

さて、分散共分散行列のモデル化ができたところで、ここまでをRで実装しておきます。

carroll <- function(r\_dat,FLG){  
   
 library(rmgarch)  
   
 #1. define variables  
 N <- NCOL(r\_dat) # the number of assets  
  
 #2. estimate covariance matrix  
 basic\_garch = ugarchspec(mean.model = list(armaOrder = c(0, 0),include.mean=TRUE), variance.model = list(garchOrder = c(1,1), model = 'sGARCH'), distribution.model = 'norm')  
 multi\_garch = multispec(replicate(N, basic\_garch))  
 dcc\_set = dccspec(uspec = multi\_garch, dccOrder = c(1, 1), distribution = "mvnorm",model = "DCC")  
 fit\_dcc\_garch = dccfit(dcc\_set, data = r\_dat, fit.control = list(eval.se = TRUE))  
 forecast\_dcc\_garch <- dccforecast(fit\_dcc\_garch)  
 if (FLG == "CCC"){  
 #Constant conditional correlation (CCC) model  
 D <- sigma(forecast\_dcc\_garch)  
 R\_ccc <- cor(r\_dat)  
 H <- diag(D[,,1])%\*%R\_ccc%\*%diag(D[,,1])  
 colnames(H) <- colnames(r\_dat)  
 rownames(H) <- colnames(r\_dat)  
 }  
 else{  
 #Dynamic Conditional Correlation (DCC) model  
 H <- as.matrix(rcov(forecast\_dcc\_garch)[[1]][,,1])  
 if (FLG == "DECO"){  
 #Dynamic Equicorrelation (DECO) model  
 one <- matrix(1,N,N)  
 iota <- rep(1,N)  
 Q\_dcc <- rcor(forecast\_dcc\_garch,type="Q")[[1]][,,1]  
 rho <- as.vector((N\*(N-1))^(-1)\*(t(iota)%\*%Q\_dcc%\*%iota-N))  
 R\_deco <- (1-rho)\*diag(1,N,N) + rho\*one  
 D <- sigma(forecast\_dcc\_garch)  
 H <- diag(D[,,1])%\*%R\_deco%\*%diag(D[,,1])  
 colnames(H) <- colnames(r\_dat)  
 rownames(H) <- colnames(r\_dat)  
 }  
 }  
 return(H)  
}

本来であれば、パッケージを使用するべきではないのですが、今日はエクササイズなので推計結果だけを追い求めたいと思います。GARCHについては再来週ぐらいに記事を書く予定です。 これで準備ができました。この関数にリターンデータを入れて、分散共分散行列を計算し、それを用いて最小分散ポートフォリオを計算することができるようになりました。

### テスト用データの収集

データは以下の記事を参考にしました。

[<https://www.r-bloggers.com/introduction-to-asset-allocation/:embed:cite>]

使用したのは、以下のインデックスに連動するETF(iShares)の基準価額データです。

①S&P500 ②NASDAQ100 ③MSCI Emerging Markets ④Russell 2000 ⑤MSCI EAFE ⑥US 20 Year Treasury(the Barclays Capital 20+ Year Treasury Index) ⑦U.S. Real Estate(the Dow Jones US Real Estate Index) ⑧gold bullion market

まず、データ集めです。

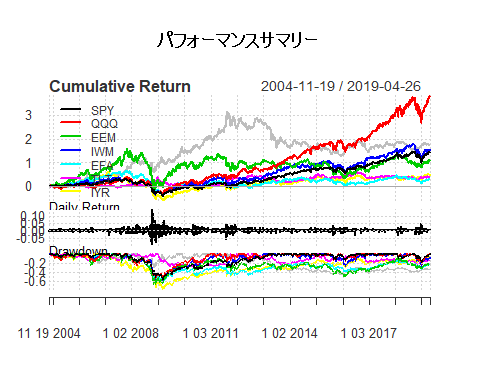
#install.packages("PerformanceAnalytics")  
library(PerformanceAnalytics)

## Warning: package 'PerformanceAnalytics' was built under R version 3.5.2

##   
## Attaching package: 'PerformanceAnalytics'

## The following object is masked from 'package:graphics':  
##   
## legend

PerformanceAnalytics::charts.PerformanceSummary(hist.returns, main = "パフォーマンスサマリー")



次に、バックテストのコーディングを行います。一気にコードを公開します。