package: rrv 0.2.0

Random Return Variables (Rough Draft)

Charlotte Maia

November 8, 2010

This vignette introduces the rrv package, a package for modelling portfolio returns as random variables. There's a strong emphasis on modelling portfolios as functions of weight, and using empirical distributions.

Introduction

This package is largely inspired by the early work of Markowitz (1952, 1959). Markowitz considered the characteristics of portfolio returns, where both portfolio returns and individual security returns are regarded as random variables. He emphasized selecting efficient sets of portfolios, giving a good compromise between expected return and variance, under the assumption that security returns are often dependent. He then creates geometric representations of portfolios (especially portfolios with three securities), using isomean lines and isovariance curves. Here, we follow from Markowitz, however with a few key differences:

1. We shall treat and emphasize portfolios as functions of weight.
2. Following the notion that portfolios are functions of weight, instead of modelling weights in a cartesian space, we shall use a triangular space.
3. We shall model historical returns using empirical distributions.
4. Whilst we shall consider variance, we shall also consider quantiles.

This package is still at a very early stage. The author is hoping to later add support for:

- Constrained optimisation.
- Modelling isoquantile curves (i.e. the quantile equivalent to isovariance curves).
- Triangular contour plots.

We shall make use of the dataset from Markowitz (1959), it gives discounted returns for nine securities over an eighteen year period.

```
> x = markowitz1959data ()
> x [c (1:2, 17:18),]
   Year Am.T. A.T. & T. U.S.S.   G.M. A.T. & Sfe   C.C.   Bdn. Frstn.   S.S.
1  1937 -0.305    -0.173 -0.318 -0.477     -0.457 -0.065 -0.319 -0.400 -0.435
2  1938  0.513     0.098  0.285  0.714      0.107  0.238  0.076  0.336  0.238
17 1953 -0.010     0.035  0.006 -0.072     -0.027  0.067  0.210 -0.084 -0.048
18 1954  0.154     0.176  0.908  0.715      0.469  0.077  0.112  0.756  0.185
```

Portfolios as Functions of Weight

Portfolio return Y, which is itself a random variable, can be regarded as the dot product of a vector of weights $\mathbf{w}$ and a vector random variable $\mathbf{X}$. Equivalently, it can also be regarded as the weighted sum of multiple (often dependent) random variables $X_1, X_2, ..., X_k$.

$$\begin{aligned} Y &= \mathbf{wX} \\ &= w_1 X_1 + w_2 X_2 + ... + w_k X_k \end{aligned}$$

We will assume that the weights always sum to one, hence it's more precise to regard portfolio return as a weighted average (rather than weighted sum). This also allows us to model weights in a triangular space (rather than a cartesian space).

Before we continue, we need to clarify two important notions, function valued functions and random variable valued functions. Perhaps the most common example of a function valued function is differentiation. When we differentiate a function, differentiation itself can be regarded as a function, say diff, that maps a function to a function. So $\mathrm{f}' = \mathrm{diff}(\mathrm{f})$. Just as functions can return functions, functions can also return random variables. Perhaps the most common example is the mean (of random variables). The mean of the elements of a vector variable random, can be regarded as function, that maps a vector random variable to a scalar random variable. So $Y = \mathrm{mean}(\mathbf{X})$.

In our case, we will construct a portfolio g, from a set of historical returns. We have a portfolio constructor $\mathrm{C_g}$, which is a function, that maps a matrix of historical returns $\mathbf{x}$ (the realised values of $\mathbf{X}$) to a portfolio. We shall regard a portfolio as a function that maps a vector of weights to portfolio return. So:

$$\begin{aligned} g &= \mathrm{C_g}(\mathbf{x}) \\ Y &= \mathrm{g}(\mathbf{w}) \end{aligned}$$

The portfolio is based on the weighted sum given earlier, and treats the random variables as constants (constant in the sense that their distributions are constant). Exactly what g, $\mathbf{X}$ and Y are, is discussed later. For now, let's take things a step further, and derive conditional parameters of Y.

Following the notion that a portfolio is a function of weight, we can also compute the expected value of portfolio return as a function of weight. This function is constructed from a portfolio, so:

$$\begin{aligned} \mathrm{f}_{\mathbb{E}} &= \mathrm{C}_{\mathbb{E}}(\mathrm{g}) \\ \mathbb{E}(Y|\mathbf{w}) &= \mathrm{f}_{\mathbb{E}}(\mathbf{w}) \end{aligned}$$

We are going to be unorthodox, and denote variance using $\mathbb{V}$ and quantiles as $\mathbb{Q}$. We compute can compute them (and almost any conditional parameter) in the same manner, so:

$$\begin{aligned} \mathrm{f}_{\mathbb{V}} &= \mathrm{C}_{\mathbb{V}}(\mathrm{g}) \\ \mathrm{f}_{\mathbb{Q}} &= \mathrm{C}_{\mathbb{Q}}(\mathrm{g}, p) \\ \mathbb{V}(Y|\mathbf{w}) &= \mathrm{f}_{\mathbb{V}}(\mathbf{w}) \\ \mathbb{Q}(Y|\mathbf{w}) &= \mathrm{f}_{\mathbb{Q}}(\mathbf{w}) \end{aligned}$$

Note that the p in the constructor for quantile return is probability, and is a number between zero and one.

Using rrv we can construct a portfolio for the two investment (or two security) case, using say, the first two securities in the dataset. We can then go on to create functions for expected return, variance (of) return, and quantile return.

```
> names (x) [2:3]
[1] "Am.T."     "A.T. & T."
```

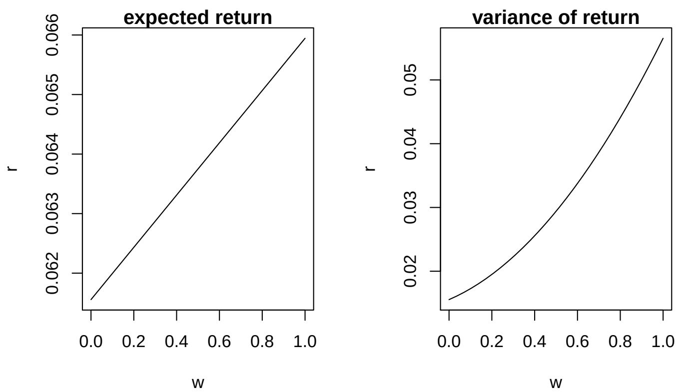
```
> g = portfolio (x [,2:3])
> fe = expectedpr (g)
> fv = variancepr (g)
> fq25 = quantilepr (g, 0.25)
> fq50 = medianpr (g)
```

If we want, we can compute the expected return of a portfolio, which contains equal weighting over both investments.

```
> fe (c (0.5, 0.5) )
[1] 0.06375
```

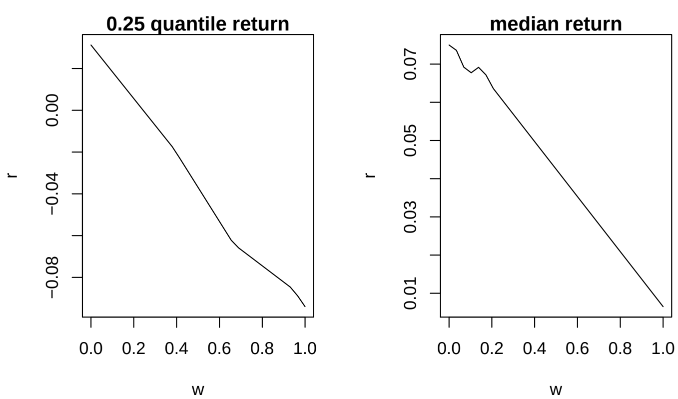
We can also plot the expected return and variance of return.

```
> par (mfrow=c (1, 2) )
> plot (fe, main="expected return")
> plot (fv, main="variance of return")
```



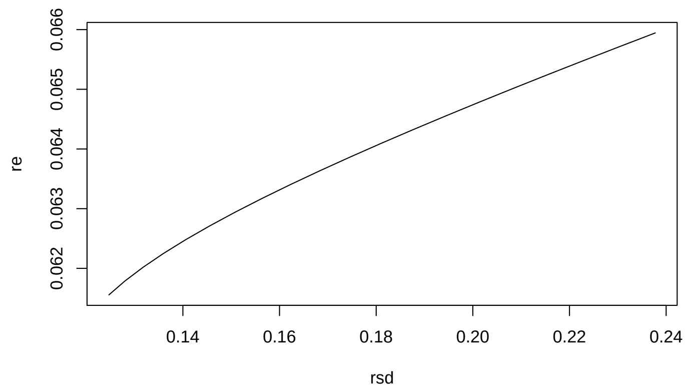
As well as, the 0.25 and 0.5 quantiles.

```
> par (mfrow=c (1, 2) )
> plot (fq25, main="0.25 quantile return")
> plot (fq50, main="median return")
```



Many textbooks consider standard deviation versus expected return. At present the implementations are not fully vectorised, so:

```
> s = seq (0, 1, length=20)
> re = rsd = numeric (20)
> fsd = sdpr (g)
> for (i in 1:20)
 {        w = c (s [i], 1 - s [i])
         re [i] = fe (w)
         rsd [i] = fsd (w)
 }
> plot (rsd, re, type="l")
```

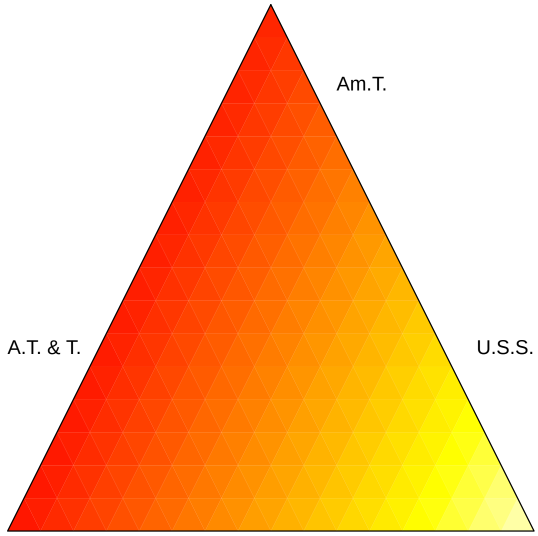


Visualising Portfolios over Triangular Spaces

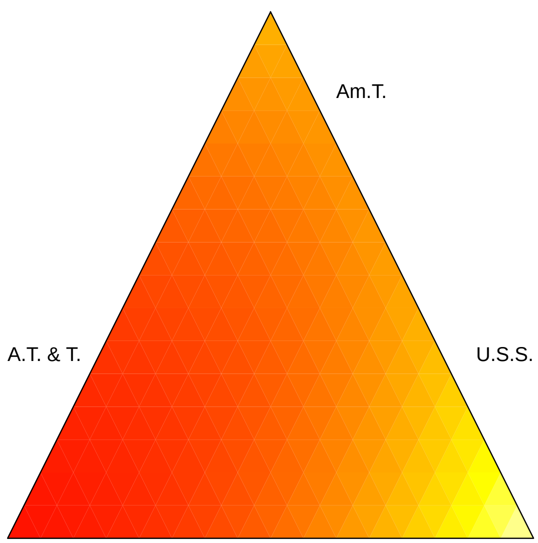
We have suggested that weight exists in a triangular space (or equivalently, that portfolios are triangular functions). This principle is most intuitive in the case of three investments.

The examples used in the previous section, works using three investments as well. Hopefully contour plots will be implemented soon. Currently, this uses heat maps, with bright colours (white and yellow) representing high values and darker colours (red) representing low values.

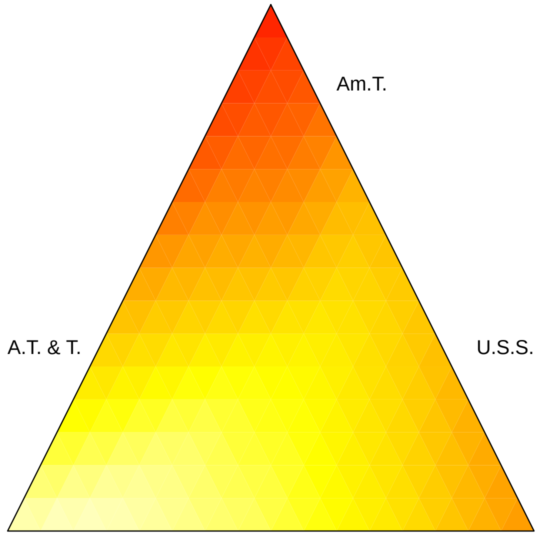
```
> names (x) [2:4]
[1] "Am.T."     "A.T. & T." "U.S.S."
> g = portfolio (x [,2:4])
> fe = expectedpr (g)
> fv = variancepr (g)
> fq25 = quantilepr (g, 0.25)
> fq50 = medianpr (g)


> plot (fe)
```



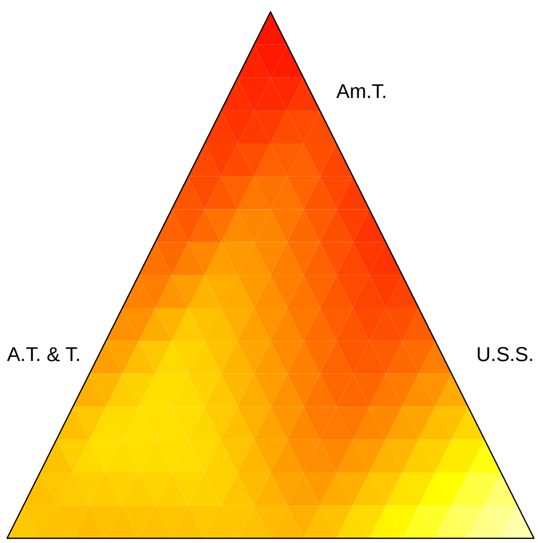

```
> plot (fv)
```




```
> plot (fq25)
```




```
> plot (fq50)
```

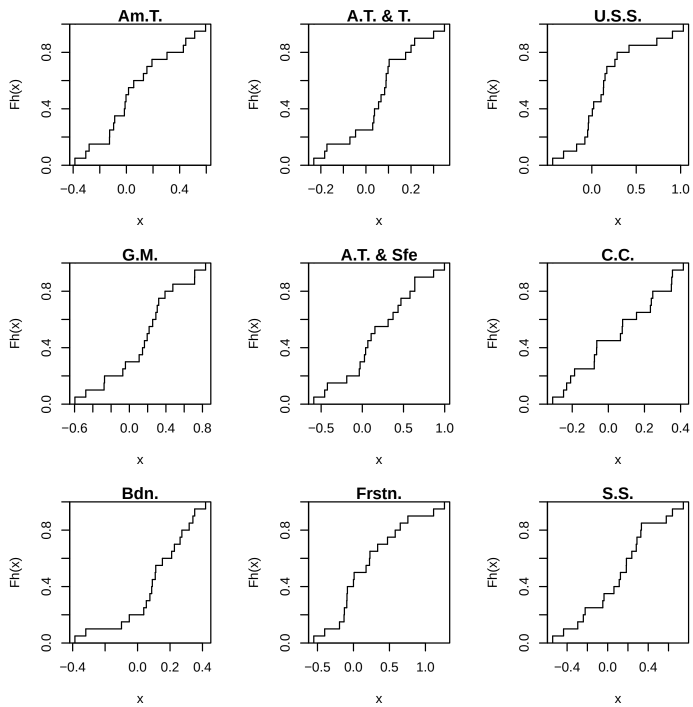


Returns with Empirical Distributions

Up to this point, we haven't really discussed what $\mathbf{X}$ and Y really are. This package assumes that random variables have an almost arbitrary distribution that can be modelled using an empirical cumulative distribution function (ECDF). In the case of vector random variables, we assume a multivariate ECDF, however at present this package does not make use of multivariate ECDFs (even though we will use the mecdf function).

ECDFs are modelled uses realised values of a random variable, and we can plot their distributions. Using all our data:

```
> par (mfrow=c (3, 3) )
> for (j in 2:10)
         plot (mecdf (x [,j], continuous=FALSE), main=names (x) [j])
```



Note that at the time of writing, there's is a minor error in the mecdf package when computing continuous ECDFs (which is the default in the univariate case).

Roughly speaking, we can compute a vector of realised values for Y, which we shall denote $\mathbf{y}$. This uses a trivial extension to the expression used at the beginning of section two (on modelling portfolios as function). Remember that $\mathbf{x}$ is a matrix (not a regular vector).

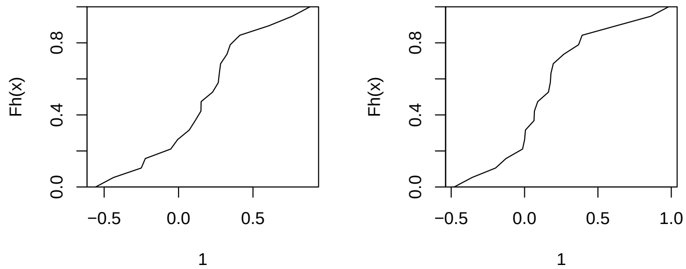
$$\mathbf{y} = \mathbf{xw}$$

In this package, we can create rrv (random return variable) objects from matrices (or any object that can be converted to a matrix). We can also create rprv objects (random portfolio return objects), using an rrv object and a vector of weights. rprv objects are what is returned by portfolio functions and are based on the expression above.

We can create a portfolio, then compute (and plot) different rprv objects.

```
> names (x) [4:5]
[1] "U.S.S." "G.M."
> g = portfolio (x [,4:5])
```

```
> par (mfrow=c (1, 2) )
> plot (g (c (0.25, 0.75) ) )
> plot (g (c (0.75, 0.25) ) )
```



Utility Functions

Note, here "utility", is utility in the programming sense.

A simple issue, which needs to be dealt with in R, is formatting numeric vectors to look like money. The can be achieved with the moneyf function (or creating a money object, see the Rd file).

```
> #some random numbers...
> x = 4e7 * rnorm (30)
> print (moneyf (x), quote=FALSE)
 [1]  32,682,242  24,351,289 -63,391,745 -34,926,118  29,513,379 -55,576,562
 [7]  56,297,302  27,638,922  53,284,265 -73,496,828 -61,129,770  37,892,672
[13] -18,003,842 -54,669,721  30,283,517   6,223,924  -5,434,243  48,268,097
[19]  14,700,474  22,208,457  -4,857,648    -127,502  39,595,924  74,589,642
[25] -51,638,108 -71,754,288  -8,327,605   2,226,115   4,658,491 -51,321,633
> print (moneyf (x, 2, "$ "), quote=FALSE)
 [1]  $ 32,682,242.39  $ 24,351,288.65 -$ 63,391,745.05 -$ 34,926,118.02
 [5]  $ 29,513,378.61 -$ 55,576,561.77  $ 56,297,302.47  $ 27,638,922.04
 [9]  $ 53,284,264.85 -$ 73,496,827.97 -$ 61,129,769.68  $ 37,892,671.89
[13] -$ 18,003,841.95 -$ 54,669,721.01  $ 30,283,517.00   $ 6,223,923.58
[17]  -$ 5,434,243.45  $ 48,268,097.47  $ 14,700,474.47  $ 22,208,457.29
[21]  -$ 4,857,648.23    -$ 127,501.65  $ 39,595,924.10  $ 74,589,641.74
[25] -$ 51,638,108.41 -$ 71,754,288.17  -$ 8,327,604.93   $ 2,226,115.28
[29]   $ 4,658,490.51 -$ 51,321,633.01
```

References

Markowitz, H.M. (1952). Journal of Finance. Portfolio Selection.
Markowitz, H.M. (1959). Cowles Foundation. Portfolio Selection Efficient Diversification of Investments.