

Dr. H. D. Vinod

Professor, Economics Department,
Fordham University, Bronx, NY

Hands-on Optimization Using the R-Software

INTRODUCTION AND MEAN VARIANCE PORTFOLIO

The special issue of this Journal is devoted to optimization. As a calculus tool, optimization methods have been well-known for over two centuries. Choice of an optimum investment portfolio has been formulated as a constrained (Lagrangian) optimization problem for many decades, with a popular version by Markowitz (1959). Vinod and Reagle (2005) describe a version of that model in matrix notation as follows.

Assume that there are n possible assets in which the investor can invest her money. Given information regarding the past performance of investments in these assets, the investor must decide how much of her funds should be allocated to each of these assets. This is a stochastic (probabilistic not deterministic) optimization problem involving beliefs about uncertain future returns of assets, irrespective of past performance. We assume that using past returns, reliable forecasts of future probability distributions of returns of all assets are available. Initially, it is convenient to assume that asset returns are Normally distributed, so that means, variances and covariances describe the entire (multivariate Normal) density, without having to worry about skewness, kurtosis or other higher moments.

For example, consider only two assets B and C. The random future returns might be denoted by random variables B and C, where the (forecasts of) mean returns are: m_b and m_c , forecast variances are: s_b^2 and s_c^2 , and forecast covariance of returns is: COV_{bc} . All are assumed to be available. The risk-adjusted returns can be easily computed if $COV_{bc} = 0$ from the two ratios: (m_b/s_b^2) and (m_c/s_c^2) . Then the solution to the optimization problem is obtained by simply ordering the risk-adjusted return ratios and choosing the top (few) assets. If asset B offers the highest ratio, the optimal choice is to invest all available (money) funds in asset B, and nothing in assets C. That is, the optimal choice allocates the funds in proportion to the following weights $w_b = 1$ and $w_c = 0$. The allocation of half the funds in each asset is stated as: $w_b = 0.5$ and $w_c = 0.5$. There is considerable generality in the solution above, since the available funds could be in any amount and any currency. The popular 'Sharpe ratio' commonly

Abstract

The aim of this paper is to bring to the attention of business professionals the power of using free R software for practical optimization problems. We discuss the derivation of the traditional portfolio selection problem showing how R can depict the mean-variance tangency result with several sophisticated extensions available. Another example is from optimization of ill-behaved functions is illustrated with Rosenbrock's banana function. Wiener-Hopf-Whittle model for dynamic stochastic optimization shows how to construct a behavioral model using regressions to study the actions of regulators.

used by investors worldwide relies on the above solution, except that the risk adjusted return is obtained by dividing by the standard deviation rather than the variance of returns.

However, zero covariances and Normality or return distributions are strong assumptions that mar general applicability of the above solution based on ranking of Sharpe type ratios discussed in Vinod and Morey (2001, 2002). Vinod and Reagle (2005) emphasize that probability distribution of returns is typically not at all Normal and why one must use a more refined measure of risk based on the 'down-side' (lower half) of probability distribution of asset returns, rather than blindly relying statisticians' symmetric measure (of scale) based on the variance (or standard deviation). In any case, the choice of the optimal allocation is correctly thought of as a choice of a vector of n weights w containing proportions which must add up to 1, since we assume that all available funds are invested in one of the n assets.

Now we write the mathematical objective function for stochastically maximizing the risk-adjusted return of the entire portfolio associated with the allocation proportions: w . Let $\mathbf{1}$ denote a $n \times 1$ vector of ones and m denote an $n \times 1$ vector of average returns. Now the expected value of the random variable P 'return of the entire portfolio' based on the allocation w is readily defined. For example, with two assets it is $E(P) = E(w_b B + w_c C) = w_b m_b + w_c m_c$. In matrix notation, this is written as: $w' m$, where the transpose of a vector is denoted by a prime. The variance of the sum of two random variables is

$$\begin{aligned} \text{Var}(w_b B + w_c C) &= E[w_b(B - m_b) + w_c(C - m_c)]^2 \dots (1) \\ &= w_b^2 s_b^2 + w_c^2 s_c^2 + 2w_b w_c \text{COV}_{ab} \\ &= w' S w \end{aligned}$$

where $w = \{w_a, w_b\}$ and S is the 2×2 variance covariance matrix with (s_a^2, s_b^2) along the diagonal and COV_{ab} as the common off-diagonal term. The variance expression in matrix notation is called a quadratic form. Of course, in general, S is an $n \times n$ matrix and w has n elements.

Relaxing the unrealistic assumption of zero covariances among all asset returns is accomplished quite simply by using an S matrix with nonzero off-diagonal elements in the quadratic form. The Lagrangian objective function then is to maximize expected returns $E(P)$ subject to some penalty for volatility $\text{Var}(P)$ involving the quadratic form (1) and the requirement that elements of w be all positive and sum to 1. We write:

$$\max_w (L) = w' m - (1/2)\gamma w' S w - \eta(w' \mathbf{1} - 1) \dots (2)$$

where γ is Lagrangian coefficient, sometimes called the coefficient of risk aversion and η is the other Lagrangian, which depends on the size of initial capital. The first order condition satisfied by the solution of the maximization problem is obtained by differentiating the expression (2) with respect to w and setting the derivative equal to zero.

By the rules of matrix algebra, $(\partial w' m / \partial w)$ is m , $(\partial w' \mathbf{1} / \partial w)$ is $\mathbf{1}$, and $(\partial w' S w / \partial w)$ is $2S w$. Hence the solution is:

$$w^* = \gamma^{-1} S^{-1} (m - \eta \mathbf{1}) \dots (3)$$

We assume that the second order condition based on the second derivative is satisfied. We discuss a variation of the solution (3) discussed in Finance texts using a mean-variance diagram (similar to Figure 1) of the portfolio theory, with mean return on the vertical axis and variance on the horizontal axis.

The graphical method makes a distinction between $(n-1)$ risky assets and one asset representing risk-free government bonds. All n assets will have their own mean and variance values, leading to a two-dimensional scatter of n points in this diagram. Since the risk-free return has zero (risk of default) variance, it represents a point on the vertical axis. When n is large, we can lump the scatter of in a few columns for the purpose of discussion. Then, the points representing highest return for each value of the variance will lie along the top part to each column. The assets lying in the lower part of a column are 'dominated,' in the sense that sensible investors will not buy them. Joining the tops of each such column traces a mean-variance 'frontier,' which will be concave (bows out).

The graphical solution is located along the set of all points along the line starting at the point representing risk-free return and ending at the point of tangency of the line with the concave top part of the frontier (See Fig. 1). A point at the exact middle of this straight line represents 0.5 weight for both the best risky asset found by the frontier tangency point and on a risk-free government bond. The Lagrangian γ representing risk aversion will help determine the relative weight on risk-free investment and the asset (s) representing the point of tangency. A risk averse investor will have a higher weight on government bonds. I have attempted to summarize certain important theoretical insights from the age-old portfolio theory. Once these fundamentals are clearly understood, it is a simple matter of using better software to consider a solution for more realistic practical situations with the help of modern R software.

This paper introduces some basics of the R software system freely available on the Internet at: (<http://www.r-project.org/>), with the Wiki help system at: (<http://wiki.r-project.org/rwiki/doku.php>). R system is ever-expanding due to contributions by a world-wide community of researchers in the form of contributed by a world-wide community of researchers in the form of contributed 'packages.' My personal notes about various tricks and pain in using R is a MS word file available on the Internet at: (<http://www.fordham.edu/economics/vinod/r-lang.doc>). A package called 'Portfolio' is useful for implementing the above theory and much more from the recent Finance literature. Diethelm Wuertz and Rmetrics Core team (<http://www.rmetrics.org>) are

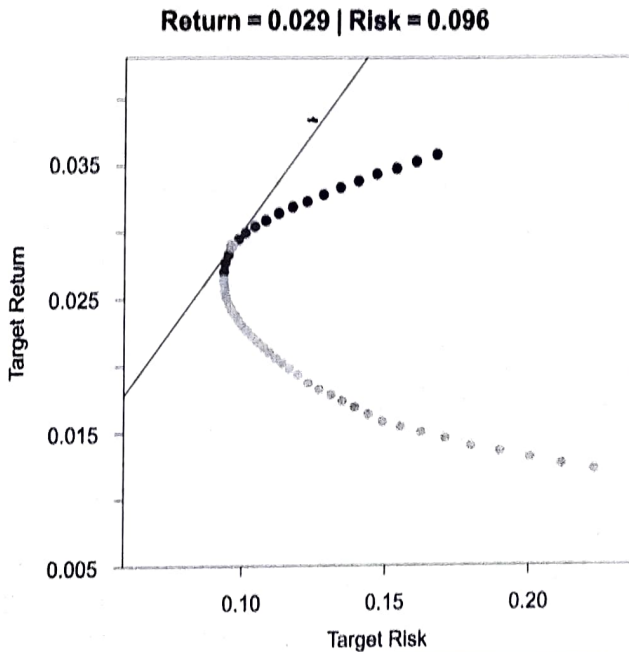
responsible for this wonderful set of dozen or so contributed packages to R all beginning with the lower case 'f'. According to the 47-page free online manual 'Portfolio.pdf' describing the package, it is currently being improved and we expect it to use 'lower partial' variance defined on the lower on the lower part of the probability of distribution (where losses occur) to measure risk, as discussed in Vinod and Reagle (2005).

The various 'Something' packages come with both US and International data useful for understanding the workings of the packages. For example, an illustrative frontier of Figure 1 can be computed by using the following snippet of R commands. Note that R is a line-by-line interpretive object-oriented language which ignores material beyond the symbol '#' on any line. This allows us to insert explanatory comments and instructions in our R code. The entire snippet can be copied and pasted into R to get the desired results.

```
library(Portfolio) #you must first load this package
#into your R session
```

```
# Load Data and Convert to time Series Object
Data = as.timeSeries(data(smallcap.ts))
#pick 4 small cap companies
#from January 1997 to December 2001
Data = Data[, c("BKE", "CG", "GYMB", "KRON")]
#Allow for unlimited Short Selling (implying negative weights)
#Constraints = "Short"
#Constraint = c("minW[1 : nAssets] = 0")
Constraints = "minW[1 : 4] = 0"
Spec = portfolioSpec() #set at default
setRiskFreeRate = 0.03
# Compute Short Selling Minimum Variance Portfolio
tangencyPortfolio(Data, Spec, Constraints)
#minvariancePortfolio(Data, Spec, Constraints)
#Portfolio Weight(s):
#   BKE   CG   GYMB  KRON
# 0.2991 0.3014 0.0000 0.3995
Frontier = portfolioFrontier(Data, Spec, Constraints)
#minvariancePortfolio(Data, Spec, Constraints)
frontierSlider(frontier) #new window will pop up
#this gives several options, click on Tangency PF
#click on slider till risk-free rate is as desired
```

Fig 1 Mean Variance Frontier Illustrated Using Portfolio Package of R



Minimizing Ill-behaved Functions It is well known that certain functions are ill-behaved and various computer algorithms for finding the minimum (or maximum) of such functions can fail to reach the correct global minimum. For example, consider Rosenbrock's banana function defined by:

$$f(x, y) = (1 - x)^2 + 100(y - x^2)^2. \quad \dots (4)$$

Its 3D plot can be seen at (http://en.wikipedia.org/wiki/Rosenbrock_function). The plot verifies that this function is ill-behaved, showing a long and narrow

valley shaped like a parabola, where its minimum lies. Since its global minimum is known to be at $x = y = 1$. This function is often used to test convergence of optimization algorithms to the global minimum. R comes with two built-in optimization functions called 'optim' and 'ConstrOptim.' It is interesting that R folks are so confident about good convergence of these functions that the manual uses the banana function for illustration. Indeed the correct minimum is achieved by both, upon supplying the analytical derivatives of the banana function (details omitted for brevity).

Wiener-Hopf-Whittle Model for stochastic dynamic optimization In the late 1940's dynamic target seeking was important for military applications and Kalman filtering won the day. Its competition was the Wiener-Hopf model, where the solution was analytic and deemed too restrictive for military applications. Sargent's foreword to Whittle (1983) notes that Whittle's frequency domain methods are useful for "deducing closed form solutions for decision rules". Vinod (1990) derives a closed form solution to a 'target seeking' minimization problem from regulatory economics.

The regulated firm (e.g., electric utility) decides its productivity, fuel costs and prices p_t , whereas external market forces eventually determine the firm's net revenue π_t . The regulator (e.g., Public Utilities Commission) sets a 'fair' rate of return π_t^* , which then is the target of a target-seeking optimization problem. In addition to attempting to reach the target, the regulator expects that prices charged by the utility will be reasonably stable over time. The variance of prices, $V_p = E(p_t - \bar{p})^2$, can be used to measure such stability.

The regulator also wants that the 'level' of prices measured by the average price \bar{p} should be as low as possible. The Lagrangian objective function is:

$$L = E(\pi_t - \pi_t^*)^2 + \mu_1 V_p + 2\mu_2 \bar{p}, \quad \dots (5)$$

where the Lagrangian coefficients satisfy $\mu_1 > 0$ and $\mu_2 \geq 0$. Note that the weights $(1, \mu_1, \mu_2)$ represent relative importance of the three terms as perceived by the regulator. The solution derived in Vinod's (1990) Appendix assumes that the regulated utility follows a linear decision rule recursion based on past profits as:

$$\pi_t = b_1 + b_2 p_t + b_3 \pi_{t-1} + \epsilon_t, \quad \dots (6)$$

where ϵ_t denotes the error term, which is absent in the normative derivation. Regarding the prices charged to customers, the normative decision rule implied by Wiener-Hopf frequency domain methods is:

$$p_t = \bar{p} + C_1(\pi_{t-1} - \bar{\pi}) + C_2(\pi_t^* - \bar{\pi}^*) \quad \dots (7)$$

where $C_1 = (\xi - b_3)/b_2$ and $C_2 = (K^2 b_2)^{-1}$ $\dots (8)$

$$\xi = (A/2) + (1/2)(A^2 - 4)^{1/2} \quad \dots (9)$$

$$A = b_3 + (1/b_2) + (b_2^2/\mu_1 b_3) \quad \dots (10)$$

$$K_2 = \mu_1 b_3 / (\xi b_2^2) \quad \dots (11)$$

where the bars denote mean values. If the error term is introduced, one can think of (6) as a regression model making b_2 and b_3 as regression coefficients of a behavioral equation (6) rather than those belonging to a normative decision rule.

Clearly, the normative decision rule is feasible only if the relative weights $(1, \mu_1, \mu_2)$ assigned by the regulator are known. Unfortunately the expressions for coefficients b_2 and b_3 from (7) to (11) are highly nonlinear and will have to be estimated from some starting values iteratively in a

nonlinear least squares type algorithm. Perhaps, the behavioral estimates from regression can provide starting values for such iterations. Any difference between such starting values and ultimate values measures the extent of the failure of human agents to optimize. Vinod (1990) rewrites (7) as a second behavioral equation upon inserting an error term u_t as:

$$(p_t - \bar{p}) = C_1(\pi_{t-1} - \bar{\pi}) + C_2(\pi_t^* - \bar{\pi}^*) + u_t, \dots (12)$$

Thus, we can estimate a system of two equations (6) and (12). Their specification based on Wiener-Hopf model assumes that both firms and regulators are 'rational' in the sense that they try to maximize their own objective functions. However, the error terms mean that the rationality is not perfect.

Equations (6) and (12) together yield regression coefficient estimates of a behavioral model used to describe how the agents including regulators actually behave. It is of interest to estimate the relative weights $[1, (V_p \mu_1), (2\bar{p} \mu_2)]$ in comparable units representing estimates of relative importance of the three terms as acted upon by the regulators. Vinod's (1990) Table 1 has annual estimates of relative weights on $(V_p \mu_1)$ representing the 'Don't Rock the Boat' motive and $(2\bar{p} \mu_2)$ as the consumer protection motive for Bell telephone data. It is interesting that between 1963 to 1970 the weight on consumer protection was near zero, which suddenly jumped in 1971 and remained somewhat high till the brea-up of Bell System in early 1982 by anti-trust litigation.

In R software the regression and any nonlinear system is readily estimated by the function 'lm' for linear regression models and 'nls' for nonlinear least squares. The package 'Nonlinear' handles most nonlinear modeling including chaos. The R package 'systemfit' fits a set of structural nonlinear equations.

FINAL REMARKS

This paper illustrates three optimization problems where R software system can make the task of finding the solutions and /or obtaining important (empirical) insights quite practical for anyone. Vinod (2008) provides R snippets as templates for extending dozens of practical problems in Econometrics. Our aim here has been to introduce the reader to the power of the R system and several new packages available free for non-commercial use to anyone with an Internet access. In particular, we have shown in Section 1 how to use 'Portfolio' package for portfolio selection problems. Section 2 describes functions 'optim' and 'ConstrOptim' for all kinds of (constrained) optimization problems including Rosenbrock's ill-behaved banana function. Section 3 proposes a way of converting difficult optimization problems into behavioral relations and using regressions to estimate them. Such estimates can throw important light

on whether regulators are properly doing their jobs consistent with their rhetoric of protecting the consumers.

REFERENCES

1. Markowitz, H. (1959) *Portfolio Selection: Efficient Diversification of Investments*. New York: J. Wiley and Sons.
2. Vinod, H.D. 1990. *Regulatory Economics: Dynamic Optimization and Control Under Multiple Objectives*, Journal of Quantitative Economics 6, 245-269.
3. Vinod, H.D. 1996. *Consumer Behavior and a target Seeking Supply Side Model for Income Determination*. Ch. 18 in M. Ahsanullah and D. Bhoj (eds). Applied Statistical Science, I Nova Science Pub. Corbondale, IL, 219-233.
4. Vinod, H.D. and M.R. Morey (2001) *A Double Sharpe Ratio*. in *Advances in Investment Analysis and Portfolio Management*, Vol. 8, Lee, Cheng Few, (ed.), New York, 2001, JAI-Elsevier Science pp.57-65.
5. Vinod, H.D. and M.R. Morey (2002) *Estimation Risk in Morningstar Fund Ratings*. Journal of Investing, Vol. 11(4), 67-65.
6. Vinod, H.D. and D. Reagle (2005), *Preparing for the Worst: Incorporating Downside Risk in Stock Market Investments*. New York: J. Wiley and Sons.
7. Vinod, H.D. (2008), *Hands-on Intermediate Econometrics Using R. Hackensack*, New Jersey: World Scientific Pub. (In Press).
8. Whittle, Peter 1983., *Prediction and Regulation by Linear Square Methods, 2nd Ed.*, Minneapolis, Minnesota: University of Minnesota Press.