ON PORTFOLIOS AND REGRESSIONS

Manuel Tarrazo, School of Business, University of San Francisco, 2130 Fulton St., San Francisco, CA 94117-1045, tarrazom@usfca.edu

ABSTRACT

This study examines the relationship between portfolios and regressions, which is desirable for educational, mathematical, and theoretical reasons. Educationally, understanding this relationship simplifies the teaching and learning of both procedures. Mathematically, portfolio optimization and regression systems are abstractly, algebraically, topologically, and structurally equivalent. One is obtained from the other as if modeling clay, without tears or discontinuities, and what one learns in one system can be applied to the other. We show portfolios and regressions are equivalent at a theoretical level as well. In the economic-financial context, this theoretical equivalence means that mean-variance, efficient portfolios are in fact optimal predictors, which is necessary for arbitrage-based investment valuation and for the study of arbitrage-based market adjustment. We use linear algebra and study the characteristics of Lagrange methods to make our point. We also provide specialized procedures to facilitate portfolio optimizations.

INTRODUCTION

Mathematically, linear regressions and portfolio optimizations share the same objects —vectors and matrices in the $\Re^k$ space of real numbers— and both procedures optimize real-valued, quadratic functions in a given coordinate system. Mathematically, portfolio optimizations and linear regressions are equivalent. The significance of this equivalency goes well beyond optimizing a quadratic function. For example, [1] have shown that excluding short sales maximizes the R-squared of the portfolio and, therefore, the exposure of the portfolio to the index; shorting securities lowers the R-squared and the systematic exposure. This makes sense; in the context of the single index model a long-only portfolio expresses confidence in market growth, while shorting securities implies the contrary. Additionally, there are many important econometric topics such as errors in variables, heteroscedasticity, stochastic variables, that offer promise for portfolio modeling. [2] and the first part of [3] focuses on portfolios and regressions. Our study unifies the perspectives and findings of these articles and provides additional clarifying material. Throughout the paper, we use a simple, numerical example readers can reproduce with readily available spreadsheet software. In the first section, we primarily clarify the procedures in [3] study. In the second section of the study, we present our own analysis, which uses coordinate systems. Regressions and portfolio optimizations are equivalent mathematically because they provide the same set of homogeneous coordinates, among other reasons. A simple transformation links regression estimates (Cartesian coordinates) to optimal portfolio weights (homogenous, barycentric). Our study of coordinate systems extends the literature on portfolio optimization. It also clarifies key issues in modern financial research such as the role of the risk free rate and the intercept in portfolio optimizations and regressions, and the numerical treatment of arbitrage. Optimal weights play a double role in portfolio optimization. The first role is the well-known one of wealth allocation ratios. The second, less known one, is that of signposts, markers –that is, a positioning system investors use to appraise the field, get oriented, and trade.

PORTFOLIOS AND REGRESSIONS

Portfolio optimizations are calculated by finding those optimal weights, wi* = {w1, …, wk}, that minimize portfolio variance, for a given return. In addition, the sum of optimal weights must add up to one. The statistical indicators (means, variances, and covariances) are calculated using stock returns. It is well-known that a solution allowing for short sales (negative weights) can be obtained using Lagrange's optimization set up: L = -σp2 -λ1 (∑ wiri – rp) -λ2 (∑ wi – 1), using obvious notation. However, quadratic mathematical programming is the tool of choice to calculate optimal portfolios because it provides solutions for both the nonnegative weights case (no short-sales allowed), and the short sales allowed case. The no short sales solution could also be obtained without mathematical programming using a variable reduction method described in [4]. Britten-Jones, [3], show that optimal portfolio weights for the tangent portfolio, which is the portfolio with the largest return-to-standard deviation ratio, can be calculated using ordinary regressions in the context of a no-arbitrage model. Britten-Jones' procedure amounts to running the set of stock returns on a vector of ones (with Tx1, T =dimension equal to the number of observations) in a regression without an intercept. Then, the optimal portfolio weights are calculated by dividing each of the beta estimates by the summation of betas. The vector of residuals "inherits" the numerical values of the optimal portfolio. This is an example of EXCEL's output for Britten-Jones procedure (Exhibit III in the study):

SUMMARY OUTPUT

Regression Statistics	
Multiple R	65535
R Square	-4.30659E+33
Adjusted R Squ	-4.4577E+33
Standard Error	0.965027363
Observations	60

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	-53.08283523	-17.6943	-19	#NUM!
Residual	57	53.08283523	0.93128		
Total	60	0			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0	#N/A	#N/A	#N/A	#N/A	#N/A
X1	-1.293799509	0.984038484	-1.31479	0.1938475	-3.264303428	0.67670441
X2	5.12006439	2.061942512	2.48313	0.0159899	0.991093962	9.24903482
X3	1.093151833	1.067034031	1.02448	0.3099383	-1.043547873	3.22985154

Britten-Jones's contribution is important for many reasons. For example, it shows the relationship among different optimization procedures, it provides an easy way to obtain portfolio weights, and it illustrates why optimal portfolios must also be optimal predictors (under a linear rule). Very importantly, the regression output provides indicators of reliability for portfolio weights (standard errors, t-ratios, and p-values), which is information that was severely lacking in portfolio analysis.

In our study, we first note that using 1/T instead of 1 in Britten-Jones' regressions make the vector of mean returns appear in the optimal regression, the X'y in $b^* = (X'X)^{-1} X'y$. This means that the only critical difference between portfolio optimizations and regressions is that the former employ the variance-covariance matrix (second moment about the mean), A, in wi* = A^{-1} R, where R is the vector of mean returns, while the regression optimization employs the matrix of second moments about the origin (X'X, where X is the T-by-k matrix of stock returns. Note that optimal portfolio weights for the tangent portfolio can also be calculated by rebalancing, see [5].

COORDINATE SYSTEMS AND ARBITRAGE

The ratios between between scaled and nonscaled optimal weights within each procedure are the same. As we know, the solutions in the regressions and the portfolio optimizations depend on ratios, not on absolute values. There is a proportionality factor in the relationship between weights across procedures. These are properties of homogenous systems, which also implies that we can regard regression estimates and portfolio weights as the coordinates of each system. Note further that if we use 1/T = 1/60, T being the number of observations in our sample, as Britten-Jones' regressand constant, we find that both systems share the same incidence points, that is, the same vector of constants c in each of their simultaneous equations systems, A x = c.

There is a relationship between homogeneous coordinates and homogeneous equation systems. Let A x = c, where —to simplify matters and to not get into generalized inverse issues— A is a square, non-singular matrix. We can build its homogenous counterpart: A x – c = 0. This homogenous system has a nonzero, or nontrivial solution, if a system of a higher dimension has a solution. Bring in another variable, say lambda, and the solutions for x in terms of lambda are obtained. Barycentric coordinates are those in which the extra-variable is set to one. This type of coordinate system appears naturally in problems where the solutions are restricted to some maximum value (e.g., the initial wealth in the portfolio optimization case). Barycentric coordinate systems provide a frame of reference; the solution forms a barycenter, known in other contexts, as "center of gravity" and "centroid".

Because barycentric coordinates are homogeneous coordinates, they provide a common vertex, a common reference point to the two systems we have been studying —regression and portfolio optimization. However, we can find that point because the systems share the vector of returns. It is interesting to observe that the value of Lambda matches the returns-variability ratio in each of the systems at the optimum. It is easy to include a risk free rate in the analysis —replace the columns of ri with excess returns, ri – rf. In sum, portfolio optimizations and regressions share the same homogeneous coordinates becase they are algebraically equivalent. More importantly, this equivalence shows that arbitrage may be implemented using different information sets which, nonetheless, may carry the same information.

REFERENCES

[1] Cheung, S., and Kwan, C. "A Note on Simple Criteria for Optimal Portfolio Selection." The Journal of Finance, Vol XLIII, No. 1, March 1988, 241-245.
[2] Stevens, G. "On the Inverse of the Covariance Matrix in Portfolio Analysis." The Journal of Finance, Vol. LIII, No. 5, October 1998, 1821-1827.
[3] Britten-Jones, M. "The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights." The Journal of Finance, Vol. LIV, No. 2, April 1999, 655-671.
[4] Tarrazo, M. "Mathematical Programming and Portfolio Optimizations: A Clarification." Advances in Financial Planning and Forecasting, Vol. 9, 2000, 31–55.
[5] Elton, E., and Gruber, E.. "Modern Portfolio Theory and Investment Analysis." 5th ed., Wiley, 1995.