46-972
MSCF Finance
Mini 1 – 2017

Bryan Routledge (routledge@cmu.edu)

Assignment 1
Due: 9.11.2017 – 5.30pm

Notes - This is our first assignment

- If you are having difficulty with the analysis or just have questions that pop up, feel free to contact me (email is easy). In particular, if you are spinning and uncertain as to where to start, contact me. The assignments cover material we have done and foreshadow things we will do. They are designed to help us learn and are not meant as "review" or "testing." Part of the learning includes discussions with your colleagues and me.
- A general suggestion, the earlier you start on things the better. Having a bit of time to ponder and discuss the questions makes the assignments more effective and less stressful. This is particularly true here as we are working with data. Data is always a bit messy and raw. The tiny little problems that pop up are usually easy to solve with some reflection (or an email to me). But at the last minute, the tiny little bugs just become stress.
- You may hand in the assignments individually or in groups. A group size of about three seems best. However, you are welcome to choose any size group you like (larger or smaller is fine). There is also no need to stay in the same group for each assignment. When you turn in your groups assignment via Canvas, just list all the group members on the first page. (If you do work in a group, please submit just one assignment per group with all names on.)
- The assignment is due at the start of class.
- Assignments should prepared "professionally" (no hand-written scraps of paper). See my suggestions (under assignments on the Canvas page). It

is good practice and with practice it is no amount of extra work.

- Assignments should be turned in via Canvas and in PDF. Please, no paper!

Notes - Specific to this assignment

- We are chatting about portfolio choice. In class I wave my hands and draw pictures. But, the best way to learn portfolio choice is to roll up your sleeves, fire up a spreadsheet (or matlab, R, or other) and calculate some means, variances, and Sharpe Ratios. In this assignment you will work with data (see csv and/or excel canvas link).
- You can calculate everything here using Excel if you like. That is a bit tedious but lets you see all the pieces. However, you likely want to do this in a more robust language (Python, R, Matlab,...). To help you with that coding, see the linear algebra at the back of this assignment. You are welcome to use whatever tool works best for you. But one caution – There are all sorts of packages and add-ins for excell, R, matlab, etc. that will do all these calculations like "magic." It is best if you do the assignment **without** any of those. You want to code this up so you can see how it works. Moving forward you can then use any of those packages and know what they are doing.
- When preparing your answers and analysis, give some thought to presentation. I do not need (want) to see all your calculations. Clip some pictures, add some text, or a table. While there is lots of work to do, you can hand in a relatively concise report. That is; hand in a PDF.
- Working with data is not hard (or any harder than any of the other questions we do), but it is not something that rushing on at the last minute is much harder. Try not to leave this assignment to the last minute. Feel free to email / stop by with questions (even at the last minute). Trust me; I have no moral high-ground when it comes to the last minute, only experience.
- Questions as needed. I am here!

Question 1: [*Returns and Risk - Time Series*]

Get the data files from Canvas (XLS or CSV). There are a couple of sheets in the spreadsheet. (There are also csv files there for you if you prefer matlab, R, or other.)

The data is monthly returns (percent per month, compounded continuously) for 49 different industries. The data runs from January 1975 to July of 2017. The data source is CRSP and compiled by Ken French: see his website for all the details if you like (He has detailed SIC definitions of the industries). The data we will work with are 8 of the industries.

$$\{Aero, Guns, Steel, Ships, Beer, Toys, Fin, Rtail\}$$

Data for all 49 industries is there just if you are curious. It is calculated by constructing a portfolio of firms in that industry and calculating the return as $\ln(P_t/P_{t-1})$ where P_t is the value of the portfolio in month t, including any dividends. Returns here are percent per month, compounded continuously (more later in our course on this). The file also includes the return (percent per month, compounded continuously) for the market (a very big portfolio of stocks that includes all listed stocks weighted by market value you can think of this like the S&P500) and the risk-free rate (the one month t-bill rate). I have also calculated the returns minus the risk-free rate. This is called the “excess returns” and we use it in question 2.

(a) Suppose you invested $1 in 1.1.1975 in each of these assets. Plot the (log of) value of this investment from 1975 to now.

When you calculate the time-series value, recall the data is percent per month and compounded continuously. :So, $V_0 = 1$, $V_t = V_{t-1}\exp(r_t)$. This has the nice property that $\log V_T = \sum_{t=1}^{T} r_t$.

Your plot will have 10 lines on it. The risk-free, the market, and eight industries. Do you best to make it look pretty, but do not worry If it is overly busy.

(b) Explain why the “market portfolio” lies (approximately) in the middle of all these industry portfolios.

(c) What happened in 1987? In 2000? In 2008? [Just a sentence or two for each. Know your financial history!]

Question 2: [*Returns and Risk - Portfolio Choice*]

We are now looking at mean-variance (or mean-standard-deviation) portfolios. The goal is to use the data to form a portfolio in September of 2017. What we really want to know is the current expected returns, variances, and co-variances. Since we cannot see these, we estimate them with the data. The current risk-free rate is about 1% per year (we will work with annual rates.). Since the risk-free rate has varied over the sample of data we have (note the high risk-free rate from the Volker[1] era from the 1980's), it is helpful to calculate means, variances, and co-variances from "excess return" data. That is: stock returns above the risk-free rate. Statistically, we are assuming that excess returns are stationary and independent and identically distributed. This is not a perfect assumption, but for excess returns it is a good statistical model to use.

(a) Use "Excess Return" data to estimate the Expected excess return on each industry portfolio.

- This is simple. Take an average of the monthly returns in each column
- Convert: Annual Average excess return = 12*monthly Average excess return (that is why we are working with continuously compounded returns). This is an estimate of the annual excess returns.
- To get an estimate of the current expected return (not the excess), just add the current risk-free rate of 1%.

(b) Estimate the Variance and Correlation

- calculate the sample correlations and variances
- To convert these monthly returns into annual numbers (which might be on a scale you are used to), use
  Annual Variance return = 12 * monthly Variance return
  (This uses the assumption that each observation in the data is distributed *iid*).
- Annual Correlation = monthly Correlation (remember to convert correlation to covariance when you get working with portfolios below)

[1] Paul Volker was chair of the Federal Reserve Board and ushered in the period of interest-rate targeting as a way of communicating and implementing monetary policy. Prior to Volker, the Fed targeted the harder-to-measure quantity of money as it set out maintaining the stability of the price system. In 1982, inflation rates above 10% were common.

(c) Calculate the Sharpe Ratio for each industry and market in the data. Recall: the Sharpe Ratio (below) is excess return divided by the standard deviation.

$$\frac{E[r_i] - r_f}{\sigma_i}$$

(d) Plot the Portfolio Frontier using "`Beer`" and "`Fin`". (I am not sure a mix of finance and beer is a good idea?). Since there are just two assets to consider, all you need to do is calculate the mean and standard deviation for several different portfolios using these two assets. That is, if you invest (x)% in "`Beer`" this implies $(1 - x)$% in "`Fin`".

(e) Suppose the risk-free rate is 1.0% per year, find the tangency portfolio. That is the combination of "`Beer`" and "`Fin`" that all mean-variance investors would hold.

An easy way to do this is to let solver maximize the Sharpe-Ratio. The idea is to make the slope of the capital-market-line as large as possible.

$$\max_{w_i} \frac{E[r_p] - r_f}{\sigma_p}$$

(f) Now add "`Steel`" to the mix. Plot the Portfolio Frontier using "`Beer`", "`Fin`" and "`Steel`". Since there are now three assets to choose from, you need to find the minimum variance portfolio that has a return greater than or equal to r% (you can use solver in Excel, for example). You plot the frontier by solving this problem for several different choices of r.

(g) Again, Suppose the risk-free rate is 1.0% per year. Find the "tangency" portfolio for assets using "`Beer`", "`Fin`" and "`Steel`".

(h) Why is the Sharpe Ratio of the tangency portfolio higher than in the two asset case? Did we get lucky in choosing to add "`Steel`" or is this a general result?

Question 3: [*Returns and Risk - Portfolio Choice - Many Assets*]

[This is harder since there are more assets. It is (much) more tedious to solve in excel. You need Matlab, R, etc. Let me know if you get stuck.]

(a) Calculate and plot the mean-variance frontier and the tangency portfolio using all eight of the industries. Plot it and show the tangency (max Sharpe Ratio) portfolio .

- Is the tangency portfolio sensible? Feasible?
- Why does adding assets increase the Sharpe Ratio?

(b) How sensitive are the Sharpe Ratios to the recent stock price changes? Calculate the Sharpe Ratios for the 8 industries:

- 1975 to December 2007 (missing the downturn of Sept 08)

Using this set of data, find the tangency portfolio. Compare the portfolio weights you get here with the result in part (a). In general, you will notice that the optimal weights are not very stable and are very sensitive to the data used to estimate means, variances, and co-variances.

Note: Designing portfolios that account for short-sale constraints, time-variation in correlations, and portfolios that are robust to estimation is a hard problem. Several of these topics come up in your future MSCF courses.

Estimation and statistics – Optional Reading

When using data to estimate moments of a distribution, you need a statistical model. Here, let $\tilde{r}_t$ be the return for month t and let $\tilde{R}_t$ be the return for the twelve month period (year). We are interested in estimating $E[R_t]$ and $V[R_t]$ (where you can think of these as scalars for one stock or a vector/matrix for many stocks. To use the data to estimate these quantities we need a statistical model. Here, the model we use is $\tilde{r} \sim IID(\bar{r}, \Sigma)$ – assuming the data is drawn from an identical distribution each period with constant mean and variance. For returns data, this is a sensible model and we will use it here.

We are working with continuous compounding returns. So $V_0 = 1$, $V_t = V_{t-1}\exp(r_t)$. This has the nice property that $V_T = \exp(\sum_{t=1}^{T} r_t)$ or $\log V_T = \sum_{t=1}^{T} r_t$. So $\tilde{R}_{year_1} = \tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3 + ... + \tilde{r}_{12}$. From this,

$$\begin{aligned} E[\tilde{R}_{year_1}] &= E\left[\tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3 + ... + \tilde{r}_{12}\right] \\ &= E[\tilde{r}_1] + E[\tilde{r}_2] + E[\tilde{r}_3] + ... + E[\tilde{r}_{12}] \\ &= \bar{r} + \bar{r} + \bar{r} + ... + \bar{r} \\ &= 12\bar{r} \end{aligned}$$

Note where we are using the IID assumption. Now, the same with variance

$$\begin{aligned} V[\tilde{R}_{year_1}] &= V\left[\tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3 + ... + \tilde{r}_{12}\right] \\ &= V[\tilde{r}_1] + V[\tilde{r}_2] + V[\tilde{r}_3] + ... + V[\tilde{r}_{12}] \\ &= V + V + V + ... + V \\ &= 12V \end{aligned}$$

Here you can see we are using the assumption that returns in February are not correlated with returns in March. Lastly, to convert to standard deviation, take the square root. (You can prove that with this IID assumption, monthly correlations are equal to annual correlations.)

So is the assumption that returns are IID sensible? Yes. As a first cut at the data it is reasonable. In particular, there is not much serial correlation in returns month to month. However, a more sophisticated statistical assumption would address: (1) Variance is not constant across time ("stochastic volatility"), (2) There is some serial correlation in returns ("momentum"), and (3) Risk premia or conditional expected returns are not constant (the "equity risk premium" – the excess return on all stocks in excess of bonds – is low in booms and high in recessions).

Matrix Algebra and Portfolio Optimization – Optional Reading

If you happen to know linear algebra and constrained optimization ("Kuhn Tucker conditions"), mean variance portfolio algebra is quite elegant. If you do not know matrix notation and calculus, this is just needless confusion so you can skip it. If you have questions, let me know. I am happy to chat! If you want to read more, Kerry Back's 2010 "Asset Pricing and Portfolio Choice Theory" (`http://amzn.to/QFbi69`) has a nice chapter on mean-variance portfolio optimization.

Notation: $\tilde{r}$ is a N-vector of random returns with mean R and variance matrix V. A portfolio ω is a N-vector of asset weights satisfying $\omega'\mathbf{1} = \sum_n \omega_n = 1$. (the "$\mathbf{1}$" is a vector.)

Quick review. Convince (or remind) yourself of the following. In all these, just write out what the matrix notation means in terms of summations:

- $E[\omega' R] = E[R'\omega] = \omega' R$
- $dE[R'\omega]/d\omega_n = R_n$ or the whole N-vector $dE[R'\omega]/d\omega = R$.
- $Var(R'\omega) = cov(R'\omega, R'\omega) = \omega' V \omega$.
- $cov(R'\omega, R_n) = \omega' V[0, 0, ..., 0, 1, 0...0]'$ (the one is in the n-th spot).
- $dVar(R'\omega)/d\omega_n = 2cov(R'\omega, R_n)$ or write the whole Nx1 vector as $dVar(R'\omega)/d\omega = 2V\omega$.

To solve a general mean-variance portfolio, we want to minimize the variance subject to having a particular mean return, called here μ.

$$\begin{aligned} \min_{\omega} \quad & 0.5\omega' V \omega \\ s.t. \quad & w'R \geq \mu \\ & w'\mathbf{1} = 1 \end{aligned}$$

Note the two constraints here are (1) that the mean return is (at least) μ and (2) that the portfolio weights sum to one. The Lagrangian for this constrained optimization is

$$\min_{\omega,\lambda_1,\lambda_2} \quad 0.5\omega' V\omega - \lambda_1(w'R - \mu) - \lambda_2(w'\mathbf{1} - 1)$$

(where λ_1 and λ_2 are the Lagrange multiplies on the two constraints)

- The First Order Conditions are:

  $$\begin{aligned} V\omega^* &= \lambda_1 R + \lambda_2 \mathbf{1} \\ \omega^{*\prime} R &= \mu \\ w^{*\prime}\mathbf{1} &= 1 \end{aligned}$$

  where the "$*$" indicates things are optimal. Rearrange the first one and you get

  $$\omega^* = \lambda_1 V^{-1} R + \lambda_2 V^{-1}\mathbf{1} \tag{1}$$

- Define two portfolios. They turn out to be handy

  $$\begin{aligned} \omega_R &= \frac{V^{-1}R}{\mathbf{1}' V^{-1} R} \\ \omega_1 &= \frac{V^{-1}\mathbf{1}}{\mathbf{1}' V^{-1}\mathbf{1}} \end{aligned}$$

  Verify these are portfolios. This is easy! Just check that they sum to one. $\omega_R'\mathbf{1} = 1$ and $\omega_1'\mathbf{1} = 1$. Notice that the denominator of each of these is just a scalar.

- Plug ω_R and ω_1 into equation (1)

  $$\begin{aligned} \omega^* &= \lambda_1 V^{-1} R + \lambda_2 V^{-1}\mathbf{1} \\ &= \lambda_1(\mathbf{1}' V^{-1} R)\omega_R + \lambda_2(\mathbf{1}' V^{-1}\mathbf{1})\omega_1 \\ &= \alpha\omega_R + (1-\alpha)\omega_1 \end{aligned}$$

where α is a scalar. That last step comes from the fact that ω^*, ω_R, and ω_1 are all portfolios (that is $\omega'\mathbf{1} = 1$). This implies that $\lambda_1(\mathbf{1}'V^{-1}R) + \lambda_2(\mathbf{1}'V^{-1}\mathbf{1}) = 1$

- So how do you get α? From the target mean return μ

$$\begin{aligned} \omega^{*\prime}R &= \mu \\ \alpha\omega_R'R + (1-\alpha)\omega_1'R &= \mu \\ \alpha\mu_R + (1-\alpha)\mu_1 &= \mu \end{aligned}$$

  where μ_R and μ_1 is the mean portfolio return for the portfolios ω_R and ω_1. So

$$\alpha = \frac{\mu - \mu_1}{\mu_R - \mu_1}$$

- Cool. You have now characterized the set of mean-variance efficient portfolios $\omega^*(\mu)$. That is, you now have the optimal mean-variance portfolio for a given level of return μ. The standard deviation for this portfolio is $(\omega^*(\mu)'V\omega^*(\mu))^{0.5}$. Now just plot it for a range of values for μ.