Risk and Return Attribution
Attribution analysis attempts to find the contribution to return or risk from underlying factors. This is either done Ex-Post or Ex-Ante.

Ex-Post looks at realized risk and return. I.E. How much did this stock contribute the the total return of the portfolio and realized standard deviation?

Ex-Ante looks at assumed forward distributions and finds the contributions to expected risk and return.

We do this by estimating the first derivative and using the fact:

$$f(x) \approx \sum_{i}^{n} x_i \frac{\delta f}{\delta x_i}$$

For return this is easy. The portfolio return is the sum of the weights times the arithmetic returns. The derivative of the portfolio return with respect to factor i is just the factor weight:

$$R_t = \sum_{i}^{n} r_{i,t} w_{i,t}$$

So for any given time, t, the contribution to return for factor i is $w_i r_i$.

The individual time period return and its attribution is an arithmetic return. Arithmetic returns are not summable. The total return is not the sum of individual period returns – the total attribution for a factor is not the sum of the individual period contributions. We need a way to link individual time period returns.

There are a few methods each with their own pros and cons. The most straightforward is Cariño's K.

The arithmetic total return:

$$R = \left[\prod_{t}^{T}\left(1 + R_t\right)\right] - 1$$

The geometric total return:

$$GR = ln(R + 1)$$

Let K be the ratio between the two:

$$K = \frac{GR}{R}$$

Create a vector of k_t:

$$k_t = \frac{ln\left(R_t + 1\right)}{K * R_t}$$

By construction:

$$R = \sum_{t}^{T} R_t k_t$$

The k vector provides a scaling factor for each day. Because the factor component returns are summable inside the period:

$$R = \sum_{t}^{T} k_t \sum_{i}^{n} w_{i,t} r_{i,t}$$

The total attribution of return for factor i is then

$$A_i = \sum_{t}^{T} k_t w_{i,t} r_{i,t}$$

Calculating Weights Through Time
Unless we rebalance the portfolio every time period, the weights of each factor will move with market movements. We need the weight at each time period to calculate the ex-post return attribution. Luckily this is straightforward to calculate:

First grow the weight by the time period return.

$$w^*_{i,t} = w_{i,t}\left(1 + r_{i,t}\right)$$

By definition, the portfolio return at time t is

$$R_t = \left(\sum_i^n w^*_{i,t}\right) - 1$$

Normalize the updated weights to get the next period's starting weight:

$$w_{i,t+1} = \frac{w^*_{i,t}}{R_t + 1}$$

This gives us a way to calculate the individual portfolio returns at each time period while calculating the individual time period weights.

Risk Attribution
In this section we will look at risk as defined as portfolio volatility.

$$\sigma_p = \sqrt{w'\Sigma w}$$

Given the formula at the start of the notes:

$$f(x) \approx \sum_i^n x_i \frac{\delta f}{\delta x_i}$$

We need to find the derivative of volatility with respect to w. You are welcome to solve this by hand. I'll just give it to you:

$$\frac{\delta\sigma_p}{\delta w} = \frac{\Sigma w}{\sqrt{w'\Sigma w}}$$

This is the gradient. For an individual w_i take the i^{th} row.

$$\frac{\delta\sigma_p}{\delta w_i} = \frac{1}{\sigma_p}\sum_j^n \left(w_j * \Sigma_{i,j}\right)$$

So the vector of component volatilities is:

$$CSD = w\#\frac{\Sigma w}{\sqrt{w'\Sigma w}}$$

Where # is the element wise multiplication operator

This gives us a way to do Ex-Ante risk attribution, Ex-Post we need to recognize something about the formula for an individual derivative.

$$\frac{\delta\sigma_p}{\delta w_i} = \frac{1}{\sigma_p}\sum_j^n \left(w_j * \Sigma_{i,j}\right)$$

The summation is the definition of:

$$\sum_j^n \left(w_j * \Sigma_{i,j}\right) = cov(x_i, portfolio)$$

Therefore:

$$\frac{\delta\sigma_p}{\delta w_i} = \frac{cov(x_i,p)}{\sigma_p} \Rightarrow RA_i = \frac{w_i cov(x_i,p)}{\sigma_p} = \frac{cov(w_i x_i,p)}{\sigma_p}$$

Remember the formula for the OLS β

$$\beta = \frac{cov(x,y)}{\sigma_y^2}$$

So the risk attribution is the regression coefficient of the weighted returns regressed on the portfolio return, multiplied by the portfolio standard deviation:

$$x_i = c + \beta_{i,p} p + \epsilon$$

$$RA_i = \frac{cov(w_i x_i, p)}{\sigma_p} = \sigma_p \beta_{i,p}$$

Problem 1:
Part 1 – Calculate the maximum SR portfolio using the method from last week's homework for the following stocks: AAPL, MSFT, BRK-B, CSCO, and JNJ.

Use the returns from the end of the history (1-14) until the end of February.

Calculate the Ex-Post Return Attribution for each Stock.

(See code)
Optimal Weights

Stock	Weight
AAPL	10.076%
MSFT	20.951%
BRK-B	43.839%
CSCO	8.118%
JNJ	17.015%

Return Attribution

	AAPl	MSFT	BRK-B	CSCO	JNJ	Portfolio
Total Return	-4.47%	-3.48%	-0.83%	-9.11%	-1.32%	-2.51%
Return Attribution	-0.46%	-0.71%	-0.37%	-0.75%	-0.22%	-2.51%

Part 2 – Using the same data as part 1, add the risk attribution for each stock:

(see code)

Attribution Table:

	AAPl	MSFT	BRK-B	CSCO	JNJ	Portfolio
Total Return	-4.47%	-3.48%	-0.83%	-9.11%	-1.32%	-2.51%
Return Attribution	-0.46%	-0.71%	-0.37%	-0.75%	-0.22%	-2.51%
Vol Attribution	0.16%	0.32%	0.48%	0.09%	0.14%	1.19%

Attribution to a Factor
In the above sections, we attributed risk and return to the underlying position. If we used something like the Fama French factors to describe our portfolio, how would we do the attribution?

Assume you have m factors and n assets

$$x_i = \alpha + \beta_i F + \epsilon$$

$$\beta_i F = \sum_j^m \beta_{i,j} F_j$$

Because of the α & ϵ in the regression, the attribution to the factors will not be perfect (unless $\alpha = 0$ and $\sigma_e = 0$). We have to account for this error.

$$e_i = \alpha_i + \epsilon_i$$

The portfolio can be written as:

$$p = \sum_i^n w_i \left(\beta_i F + e_i\right)$$

$$p = \sum_i^n w_i \sum_j^m \beta_{i,j} F_j + \sum_i^n w_i e_i$$

Depending on the requirements, we can calculate the error attribution to each asset or to take the sum and treat the error as a single factor for the portfolio. Often in Ex-Post analysis, we call this total the portfolio Alpha.

$$\alpha_p = \sum_i^n w_i e_i$$

We can simplify the problem creating a new factor weight and rearranging the portfolio equation:

$$w_j = \sum_{i}^{n} w_i \beta_{i,j}$$

$$p = \sum_{j}^{m} w_j F_j + \alpha_p$$

The first derivatives of this equation are straightforward and we can proceed as we did before.

1. Calculate factor weights through time
2. Calculate portfolio return through time
3. Calculate alpha through time
4. Calculate the Cariño K scaling factors
5. Perform the Return Attribution
6. Use the factor weights and returns to run the regression to find the Risk Attribution

If you want to attribute the residual risk and return to the underlying assets, expand out the α_p equation.

$$p = \sum_{j}^{m} w_j F_j + \sum_{i}^{n} w_i e_i$$

Either way, the math is the same.

Problem 2:
Using the same data as Problem 1, attribute realized risk and return to the Fama French 3+Momentum model. Report the residual total as Portfolio Alpha.

(see code)

	Mkt-RF	SMB	HML	Mom	Alpha	Portfolio
Total Return	-5.8%	-0.11%	2.03%	0.82%	1.78%	-2.51%
Return Attribution	-4.45%	0	0.29%	-0.08%	1.77%	-2.51%
Vol Attribution	1.00%	-0.01%	-0.07%	0.01%	0.27%	1.19%

Risk Budgeting and Risk Parity
Constructing a portfolio based on risk and return as described in last week's notes usually creates a non-optimal portfolio.

Richard Michaud discusses this in his 1989 paper, "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?" Michaud shows that small changes to correlations, volatilities, and expected values can change the portfolio dramatically. This has led to the often stated:

> "*Mean Variance Optimization is Error Maximizing.*"

Or as Yogi Berra said:

> "*It's tough to make predictions, especially about the future.*"

In practice, we make long term portfolio plans using estimated expected returns and covariance. Research has shown that in the long run risk premia are fairly stable.

For shorter term problems we are often less confident about the expected return – return volatility dwarfs the expected return. However, we are more confident about volatilities and to some extent correlations.

Ex-Ante risk contributions, expressed as a % of the total portfolio risk, give us an idea of how much risk we are giving to each position.

If we do not have a good view on expected return, we can instead control how much risk each of our positions uses. This is Risk Budgeting.

The Ex-Ante Risk Budget of the Mean Variance portfolio from Problem 1 (see code).

AAPL	MSFT	BRK-B	CSCO	JNJ
11.8%	29.5%	39.7%	9.15%	9.8%

Here, we have AAPL, CSCO, and JNJ with roughly the same risk budget. MSFT has 3x their assigned risk, and BRK-B has 4x.

What if we were to set these risk budgets equal? This is called Risk Parity.

You can solve for risk parity by setting up a least squares problem. Unfortunately, because we are dealing with the inverse of portfolio volatility, which also involves a square root, this has to be solved with a non-linear optimization.

$$CSD = w \# \frac{\Sigma w}{\sqrt{w'\Sigma w}}$$

$$SSE = \sum_{i}^{n} \left(CSD_i - \overline{CSD}\right)^2$$

$$min\, SSE$$

$$s.t. \sum_{i}^{n} w_i = 1$$

For the covariance matrix used in Problem 1, here are the weights under risk parity (see code)

Stock	Weight	Risk Budget	σ
AAPL	15.4%	20%	25.6%
MSFT	14.3%	20%	25.5%
BRK-B	26.5%	20%	15.2%
CSCO	15.1%	20%	23.4%
JNJ	28.7%	20%	14.9%

Lower risk stocks are given larger weights, roughly in proportion to the inverse of their individual risk (σ).

We can approximate Risk Parity by using the normalized inverse volatility. These are good starting values for your optimization if you have trouble with convergence.

$$w_i = \frac{1}{\sigma_i} * \frac{1}{\sum_j^n \frac{1}{\sigma_j}}$$

If all the correlations are equal, the Risk Parity portfolio is the inverse volatility portfolio.

If all correlations are equal and all Sharpe Ratios are equal, then the Risk Parity portfolio is also the Super Efficient, Maximum Sharpe Ratio, portfolio.

If we think that long term Sharpe Ratios between assets are equal, then the inverse volatility portfolio is an approximation of the super efficient portfolio. This is an estimate that does not rely on estimation of means or correlations, removing a source of error.

Risk Parity with NonEqual Risk Budgets

Often portfolio managers have an opinion on which assets will outperform other assets in their portfolio. One way to express this view is to give high conviction positions a larger risk weight and low conviction positions a smaller risk weight.

We can add this to our optimization by introducing a vector of relative risk budget.

$$b = \left[b_1, b_2, \dots, b_n\right]$$

If asset i should have 2 times the risk budget of all the others, then

$$b_2 = 2b_i \,\forall\, i \neq 2$$

Calculate the CSD as before

$$CSD = w\# \frac{\Sigma w}{\sqrt{w'\Sigma w}}$$

Modify CSD by element wise division by the risk budget

$$CSD^* = \left[\frac{CSD_1}{b_1}, \frac{CSD_2}{b_2}, \ldots , \frac{CSD_n}{b_n} \right]$$

$$SSE = \sum_i^n \left(CSD_i^* - \overline{CSD^*} \right)^2$$

$$min\, SSE$$

$$s.t. \sum_i^n w_i = 1$$

Example, using the data from before. Assume we have high conviction in MSFT and very little confidence in JNJ. We want MSFT to have 2 risk shares and we want JNJ to only have 50% of a risk share.

Stock	Weight	Risk Budget	B	RP Weight
AAPL	13.8%	18.2%	1	15.4%
MSFT	23.2%	36.4%	2	14.3%
BRK-B	29.1%	18.2%	1	26.5%
CSCO	14.7%	18.2%	1	15.1%
JNJ	19.3%	9.1%	0.5	28.7%

Note that this is not purely a rescaling of the prior Risk Parity weights. You can see that even though the BRK-B risk budget went down (18.2% vs 20% in RP), the portfolio weight increased.

Correlations matter.

	AAPL	MSFT	BRK-B	CSCO	JNJ
AAPL	1.00	0.61	0.00	0.20	(0.08)
MSFT	0.61	1.00	(0.04)	0.38	(0.04)
BRK-B	0.00	(0.04)	1.00	0.22	0.40
CSCO	0.20	0.38	0.22	1.00	0.17
JNJ	(0.08)	(0.04)	0.40	0.17	1.00

BRK-B and JNJ were offsetting risk from MSFT, lowering JNJ weight caused the weight of BRK-B to increase to offset the risk from the additional MSFT weight.

NonNormal Risk Parity

We have spent a lot of time discussing how returns are not normally distributed. The above risk parity portfolios all assumed multi-normal normality and used the volatility as the risk measure.

In the non-normal case, we need a more robust risk measure. We have used VaR and ES in this class. VaR is not a coherent risk measure. One of the implications of this is that the VaR surface is non-convex – optimizations using VaR may not easily converge and can converge to local optimas that are not the global optima.

Luckily ES is a coherent risk measure and convex.

We will use the same logic as before:

$$ES(w) \approx \sum_{i}^{n} w_i \frac{\delta ES}{\delta w_i}$$

However, ES does not have a closed form solution if the distribution of the assets is unknown. We can, however, estimate this using finite differences and our simulated returns, R.

$$\frac{\delta ES}{\delta w_i} \approx \frac{ES\left(\left[w_1, w_2, \dots, w_i + e, \dots, w_n\right] \mid R\right) - ES(w \mid R)}{e}$$

From here, the optimization problem is the same, just the calculation of the component risk is changed

$$CES = w\# \frac{\delta ES}{\delta w}$$

$$SSE = \sum_{i}^{n} \left(CES_i - \overline{CES}\right)^2$$

$$min\ SSE$$

$$s.t.\ \sum_{i}^{n} w_i = 1$$

You can update the above with the risk budget as before.

$$CES^* = \left[\frac{CES_1}{b_1}, \frac{CES_2}{b_2}, \ldots, \frac{CES_n}{b_n}\right]$$

$$SSE = \sum_{i}^{n} \left(CES_i^* - \overline{CES^*}\right)^2$$

Problem 3:
Using the same data as Problem 1 and assuming a 0 mean return, fit a t distribution to each stock return series. Simulate the system using a Gaussian Copula. Find the Risk Parity portfolio using ES as the risk measure.

(see coode)

Stock	ES RP Portfolio	Vol RP Portfolio	Fitted DF
AAPL	13.9%	15.4%	20,491
MSFT	14.0%	14.3%	4.6
BRK-B	26.8%	26.5%	21.7
CSCO	14.4%	15.1%	4.5
JNJ	30.81%	28.7%	9.4

AAPL returns, with a large DF, are effectively normally distributed in this simulation. Same with BRK-B. The 2 stocks with the lowest fitted df are both reduced in weight as their fat tails contribute more to tail risk.