American Options

American options allow you to exercise the option at any time up to maturity. This gives the owner "more options" to exercise. Remember "options have value," so we expect the price of an American Option to be greater than or equal to the equivalent European Option.

$$Value_{American} \geq Value_{European}$$

The decision to strike an option early means the investor will receive the value of the payoff function. However, remember the value of the European option is always greater than the payoff function.

$$Value_{American,t} \geq Value_{European,t} > \pi_t \quad \forall\, t > 0$$

If the value of the option is always greater than the payoff, the investor is better off selling the option than they would be to exercise the option and receive the payoff. This is generally the case.

This inequality will hold unless there are underlying cashflows on the asset. In that case, it might make sense to exercise the option in order to own the underlying asset to receive the cash flow. If the cash flow, plus the payoff is greater than the value of the option, then it makes sense to exercise early.

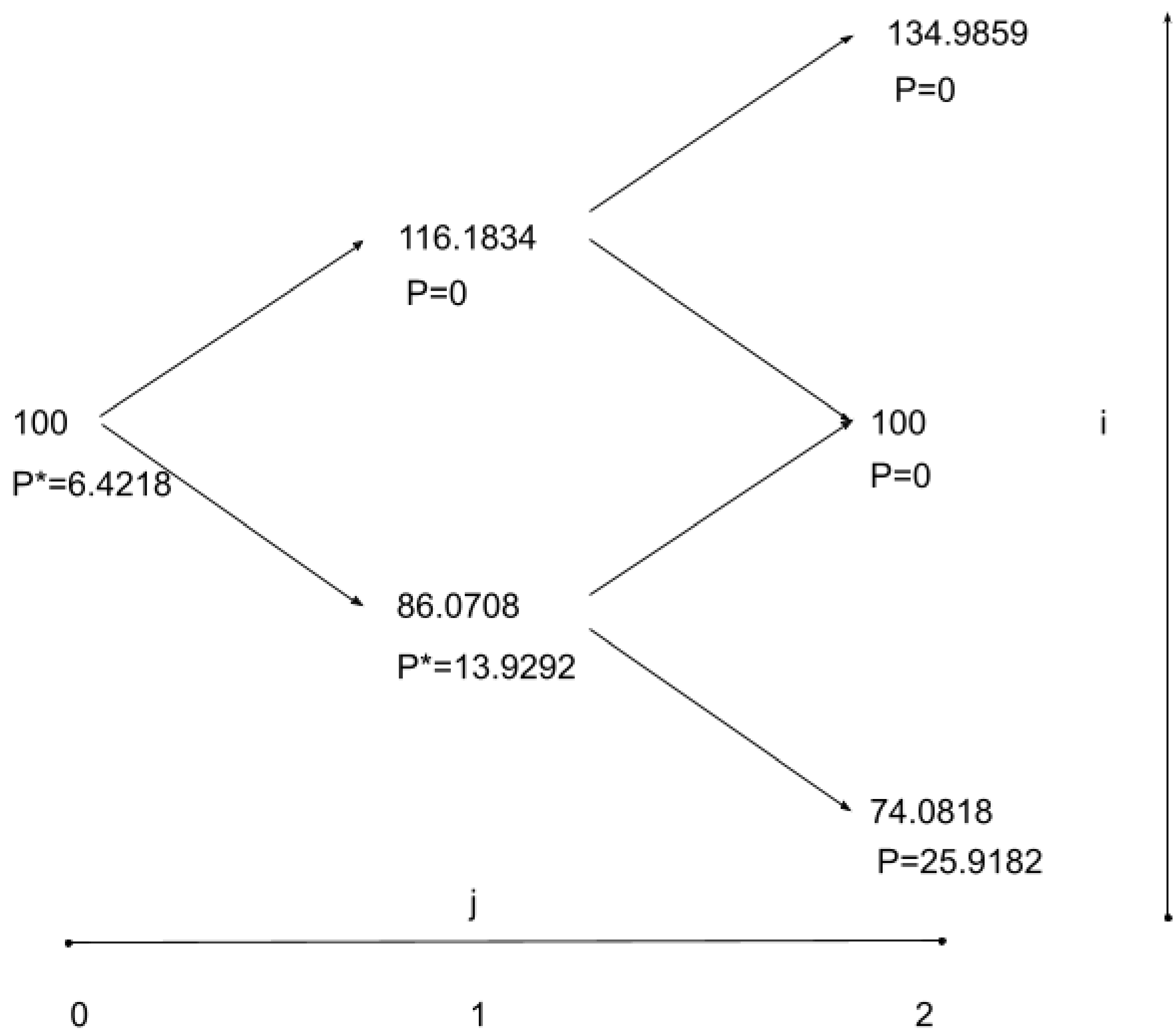
American options, as well as Bermuda options, are typically valued using a binomial tree. Here we modify the European Binomial Tree method from last week.

Steps

1. Construct the tree as before.
2. Calculate the underlying price at each node instead of just the final nodes.
3. At the maturity nodes, calculate the value of the option with the payoff function.
4. Using backward induction, check if it is optimal to exercise the option or to hold the current PV of the option from the nodes after.

$$P_{i,j} = max\left(\pi\left(Su^{i}d^{(j-i)}\right),\ e^{-r\Delta t}\left[pP_{j+i,i+1} + (1 - p)P_{j+1,i}\right]\right)$$

Example:
American Put
N=2
T=0.5
S=100
X=100
r=b=0.08
σ=0.3
$u = e^{0.3*\sqrt{0.25}} = 1.1618$, $d = e^{-0.3*\sqrt{0.25}} = 0.8607$
$p = 0.5297$



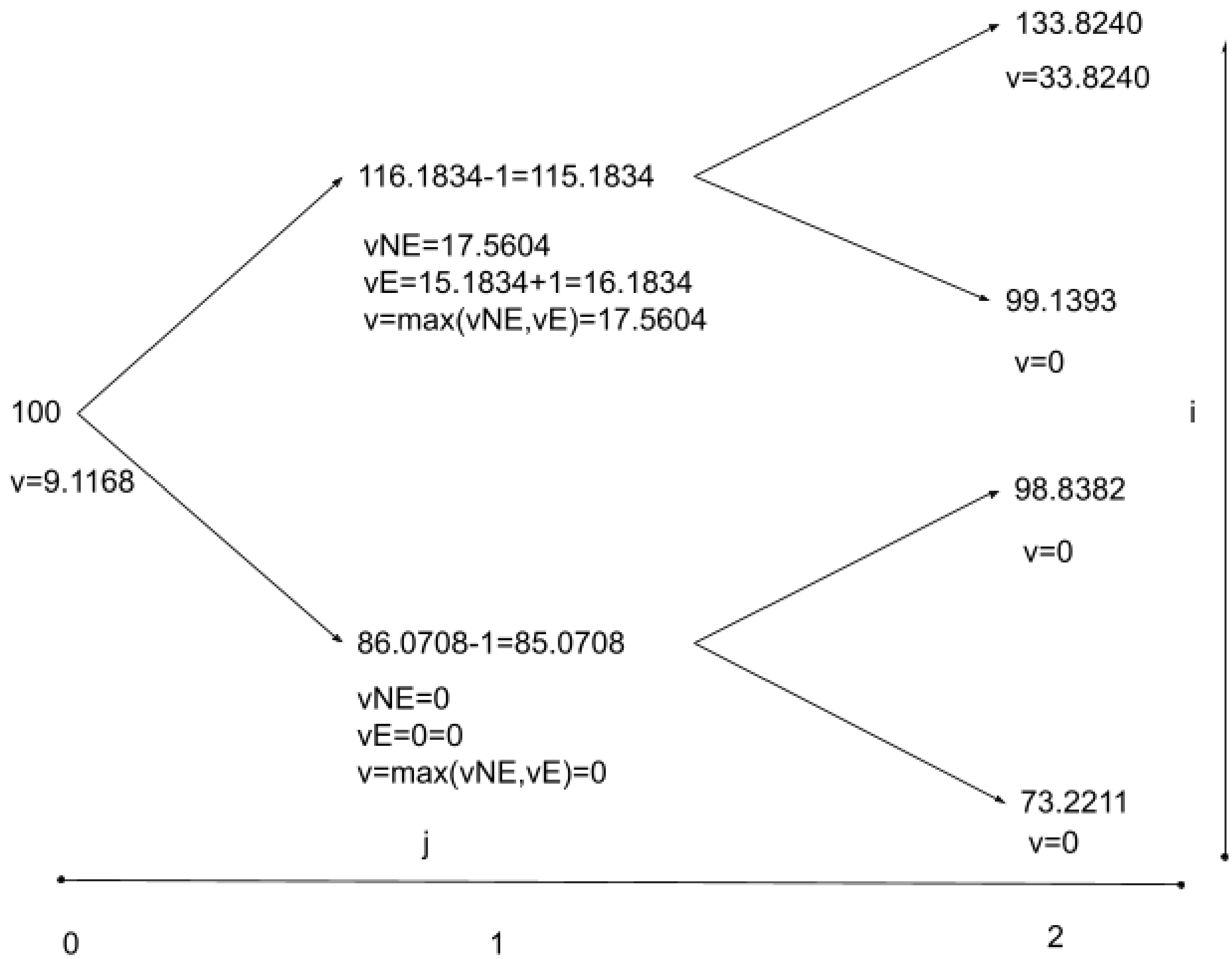
- At (j=2, i=0), the price of the put is just the payoff function.
    - At all other j=2 nodes, the value of the put is 0.
- At (j=1, i=0), the price of the put is the maximum of the payoff and the present value of the forward option prices
    - $\pi = 100 - 86.0708 = 13.9292$
    - $e^{(-0.08*0.25)}(0.5297 * 0 + (1 - 0.5297) * 25.9182 = 11.94907$
    - The option value is then 13.9292 at this node
- At (j=0, i=0) the price of the put is:
    - $\pi = 100 - 100 = 0$
    - $e^{(-0.08*0.25)}(0.5297 * 0 + (1 - 0.5297) * 13.9292 = 6.4218$
- The value of the option is 6.4218

American Options with Discrete Dividends

For index options, options on a basket of stocks, the continuously compounded dividend approach we have taken before generally works well. The basket of stocks pay their dividends at different times and they accrue to the index slowly.

However, single stocks pay dividends at a discrete time at a set value (not rate) causing a discontinuity in the share price. When a stock is paid, the price of the stock will adjust downward by the amount of the dividend. This discontinuity causes a problem in the binomial tree. Because the dividend is a set amount the percentage change in stock price is different for each node. This causes the tree to no longer recombine after the dividend is paid.

American Call
N=2, T=0.5
Dividend of 1.0 at t=0.25
S=100
X=100
r=b=0.08
σ=0.3
$u = e^{0.3*\sqrt{0.25}} = 1.1618, d = e^{-0.3*\sqrt{0.25}} = 0.8607$
$p = 0.5297$



Each Dividend paid at $j = n$ creates $n + 1$ new trees. The logic from before holds with a twist. In the case of a call exercising at $j = n$ gets the holder the value of the stock plus the dividend. In the case of a put, the holder can exercise the put, forgoing the dividend. Either way, the payoff is the same as before $\pi(S_n + D_n)$ – i.e. the payoff if the dividend didn't occur..

In practice, only call options are exercised early to receive the dividend. Hull (Chapter 15 pp 340) shows that exercise is optimal when

$$D_n > X\left(1 - e^{-r\left(T-t_n\right)}\right)$$

Where t_n is the time dividend D_n is paid.

Implementing the non-recombining tree is an exercise in recursion. This method can handle multiple dividends.

1. If there are no dividends during the life of the option, use the recombining tree method previously implemented.
2. Construct the nodes up to the first dividend payment. For each node
    a. Value No Exercise → Recursively call the non-recombining tree with the stock price minus the dividend and any remaining dividends to be paid.
    b. Value Exercise → Payoff function with the price before the dividend
3. Using backward induction, find the value at the starting node.

There are many ways to tackle the problem of discrete cash flows. Binary trees that allow for non-recombining are the most generic and allow you to apply it to many problems.

The Greeks

Option Greeks, or sensitivities, are the partial derivatives of the pricing functions with respect to one of the input variables.

These sensitivities are important because they approximate the change of the option value for changes to the inputs. Changing inputs, such as the underlying price or implied volatility cause risk to holding the option. Having a way to judge approximate changes to those inputs can allow hedging of options risk.

The most used Greeks are listed below along with their formula from the Generalized BSM formula from last week. This list is not exhaustive. People that trade options or make markets often are looking at additional cross partial derivatives to understand the risk in their portfolio.

Greek	Definition	Formula GBSM
Delta, Δ	First Derivative of Price with respect to underlying price, $\frac{\delta P}{\delta S}$	• Call: $e^{(b-r)T}\Phi(d_1)$ • Put: $e^{(b-r)T}\left(\Phi(d_1) - 1\right)$

Gamma, Γ	$\frac{\delta^2 P}{\delta S^2}$ Also called Convexity	$\frac{f(d_1)e^{(b-r)T}}{S\sigma\sqrt{T}}$, $f(x)$ is the normal PDF
Vega	$\frac{\delta P}{\delta \sigma}$	$Se^{(b-r)T}f(d_1)\sqrt{T}$
Theta, Θ	$-\frac{\delta P}{\delta T}$ Derivative is negative but often expressed as a positive number. Also called Theta Decay	• Call: $-\frac{Se^{(b-r)T}f((d_1)\sigma}{2\sqrt{T}} - (b-r)Se^{(b-r)T}\Phi(d_{1)}) - rXe^{-rT}\Phi(d_2)$ • Put: $-\frac{Se^{(b-r)T}f((d_1)\sigma}{2\sqrt{T}} + (b-r)Se^{(b-r)T}\Phi(-d_{1)}) + rXe^{-rT}\Phi(-d_2)$
Rho, ρ	$\frac{\delta P}{\delta r}$	• Call: $TXe^{-rT}\Phi(d_2)$ • Put: $-TXe^{-rT}\Phi(-d_2)$
Carry Rho	$\frac{\delta P}{\delta b}$	• Call: $TSe^{(b-r)T}\Phi(d_1)$ • Put: $-TSe^{(b-r)T}\Phi(-d_1)$

Hedging Example:
A market maker (MM) sells 100 Call options that expire in 15 trading days for $2.50. The stock price is $100.50, the strike price is $100, the risk free rate is 25bp, and the stock does not pay a dividend. The MM uses the gbsm model for pricing his options.

What is the 1 day VaR and ES for this position using a simulated normal?

The market maker hedges his position by buying stock equal to negative his delta. What is his new 1 day VaR and ES?
#####

Not all option valuation routines have closed form derivative solutions. In the case where an explicit solution is not found, a numerical method should be used. The simplest method is to use a finite difference (other methods such as [automatic differentiation] may be available and faster / more accurate).

$$f'(x) \approx \frac{f(x+\Delta)-f(x-\Delta)}{2\Delta}, \quad \lim_{\Delta \to 0} \frac{f(x+\Delta)-f(x-\Delta)}{2\Delta} = \frac{\delta f}{\delta x}$$

Given a small enough choice of Δ, the finite difference approximation for the first derivative, assuming a smooth differentiable surface, is close enough.

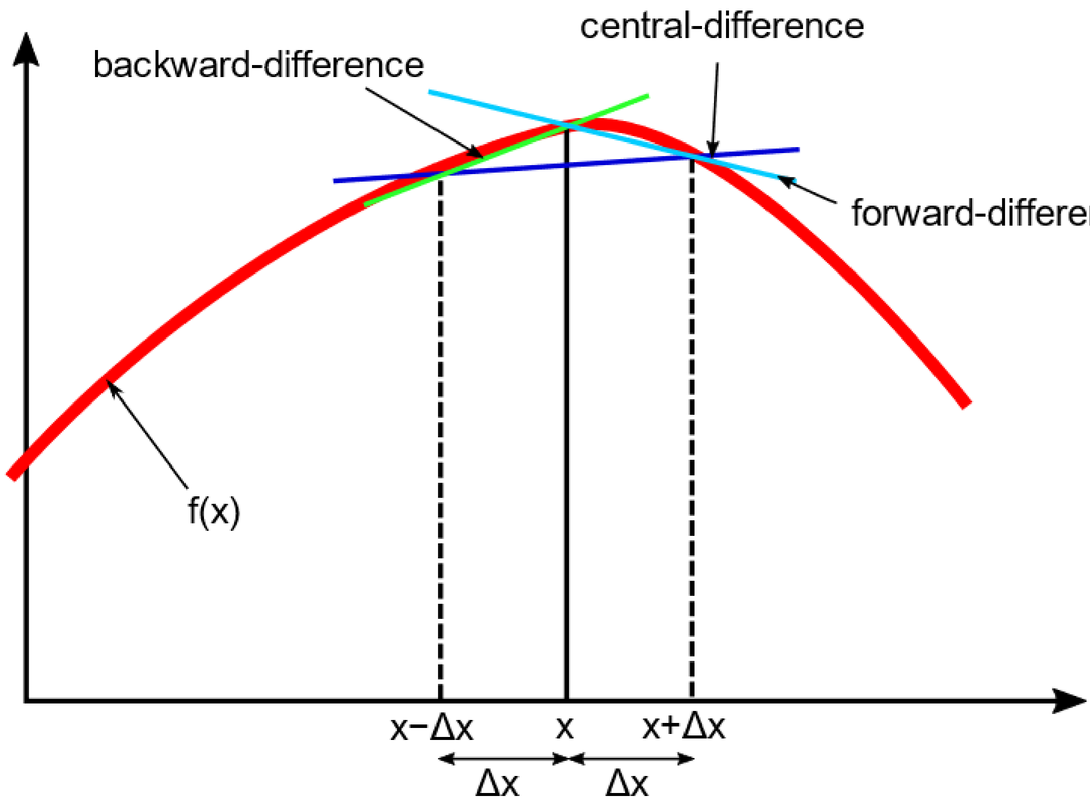
The drawback is that we have had to evaluate the function 2 times. For some valuation methods, this can be a significant computing cost burdon.

Using a backward or forward difference (instead of the central difference above) reduces the computational burden, but comes at a cost of accuracy

Forward Differencing:

$$f'(x) \approx \frac{f(x+\Delta)-f(x)}{\Delta}, \quad \lim_{\Delta \to 0} \frac{f(x+\Delta)-f(x)}{\Delta} = \frac{\delta f}{\delta x}$$

Assuming we have already evaluated the function, $f(x)$, to get the current value of the option, we only need 1 more evaluation

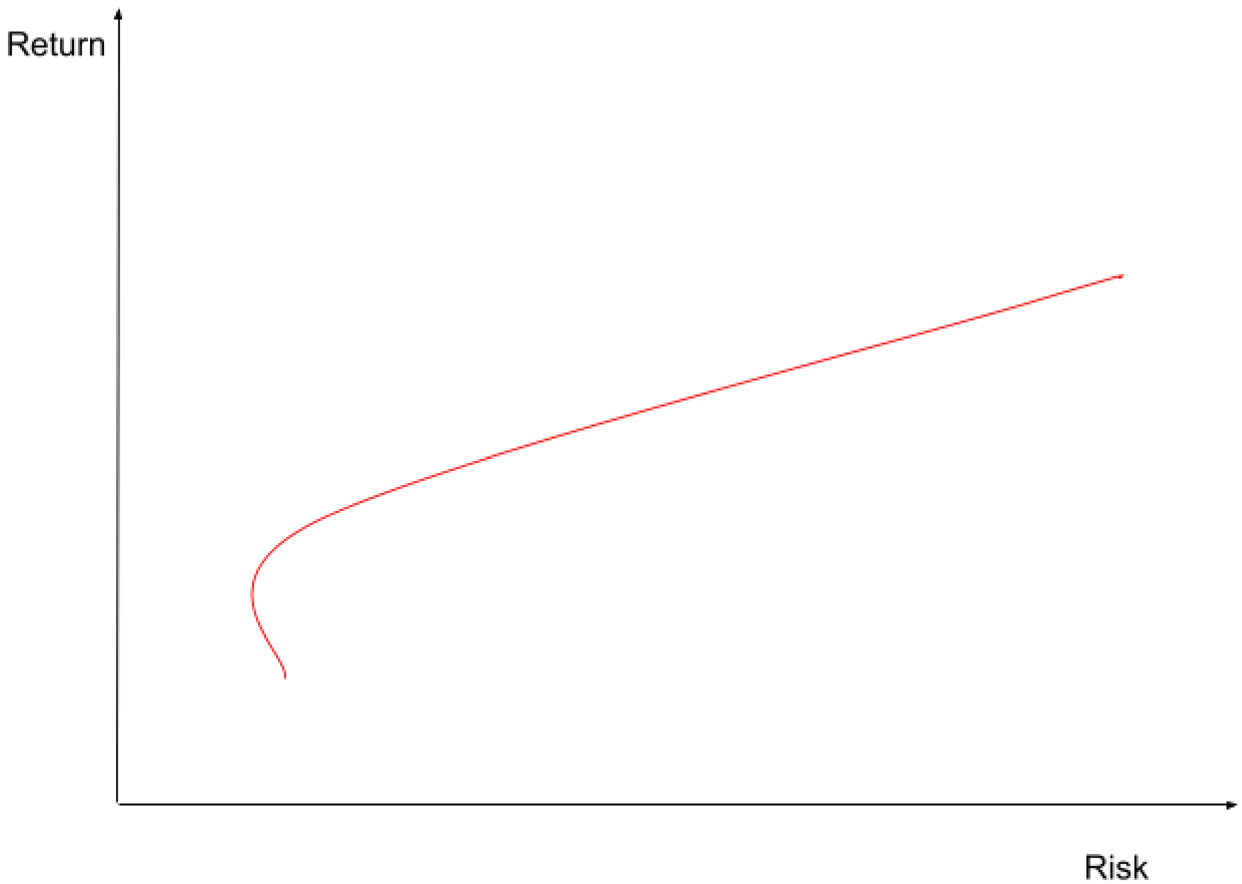


Calculating a second derivative, such as Gamma, is simply the first derivative of the first derivative.

$$f''(x) \approx \frac{\frac{f(x+\Delta)-f(x)}{\Delta} - \frac{f(x)-f(x-\Delta)}{\Delta}}{\Delta} = \frac{f(x+\Delta)+f(x-\Delta)-2f(x)}{\Delta^2}$$

Portfolio Construction
Portfolio construction is art of balancing expected return versus expected risk in a portfolio.

Generally, we trade off risk for return. Taking higher risks usually means higher expected returns.



The father of modern portfolio theory is Harry Markowitz. Professor Markowitz won the 1990 Nobel Prize in Economics for his work. The graph above is called the efficient frontier. It is the graph of the maximum return achievable by a set of investments for a given level of risk.

Markowitz assumed multivariate normality. Risk is the variance of the portfolio. The optimization problem is constructed as a quadratic problem:

$$max\ U(w) = w\mu^T + \frac{1}{2}\lambda\ *\ w\Sigma w^T$$

$$s.t.\ \sum_1^n w_i = 1$$

Where U is the utility of the portfolio's risk and return. w are the weights of the portfolio constituents. μ is the vector of expected returns. Σ is the covariance of returns. λ is a risk aversion parameter.

This setup trades off risk with return for a given investor's risk tolerance. Solving this optimization for different levels of λ will give us points along the efficient frontier.

In practice, we do not have a quadratic utility function and asking someone their "λ" value is nonsensical.

We can, however, reform the problem. Invert the graph. Recognize that each maximum return for a given level of risk is the same as the minimum risk for a given level of return.

$$min\ \sigma^2 = w\Sigma w^T$$

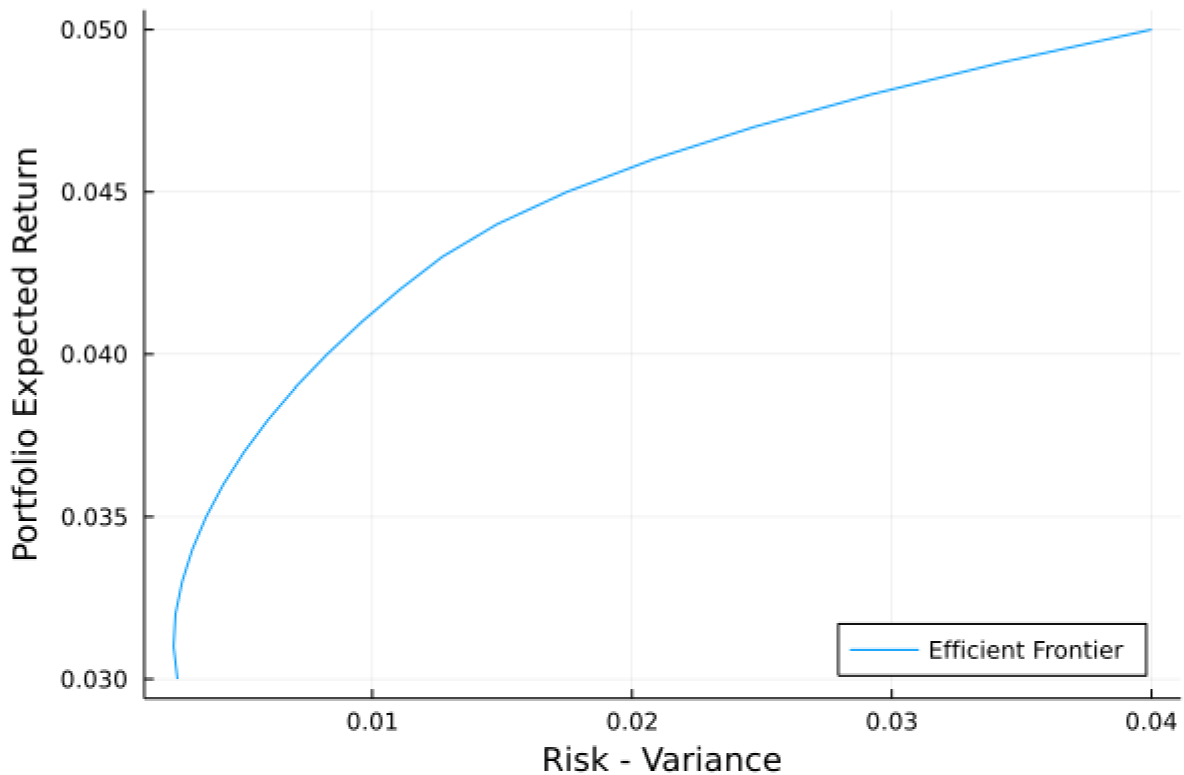
$$s.t.\ \sum_1^n w_i = 1,$$

$$w\mu^T = R$$

This is still a quadratic problem and is easily solved. Often we will put bounds on the values of the portfolio weights to either disallow shorting (negative weights) or to limit the amount short. These boundaries take the form of additional linear constraints.

Example:
Plot the efficient frontier for a portfolio of 3 assets. Do not allow negative weights.

Corr	A1	A2	A3
A1	1	0.5	0
A2	0.5	1	0.5
A3	0	0.5	1
STD	0.2	0.1	0.05
E(r)	0.05	0.04	0.03



In the mid 1960s William Sharpe and others began publishing papers on portfolio theory that led to the Capital Asset Pricing Model (CAPM). Sharpe, along with Merton Miller, shared in the 1990 Nobel Prize along with Harry Markowitz for this work. Sharpe also did work with binomial trees and their applications.

Professor Sharpe developed a metric that has been named after him, the Sharpe Ratio

$$SR = \frac{E(r_p) - r_{rf}}{\sigma_p}$$

The Sharpe Ratio compares the expected return on the portfolio in excess of the risk free rate to the risk of the portfolio, measured by standard deviation. It standardizes returns into units of risk.

Investors will always prefer assets with a higher Sharpe Ratio. Assume borrowing and lending is possible at the risk free rate. An investor can choose either investment A or investment B, but not both.

Risk Free rate: 5%
Investment A expects to return 10% per year with a 16% standard deviation.
$SR = \frac{10-5}{16} = 0.3125$
Investment B expects to return 8% per year with a 10% standard deviation.
$SR = \frac{8-5}{10} = 0.3000$

The investor only wants to have 8% risk on his portfolio.

Investment A:
Invest 50% in investment A and 50% in the risk free asset (cash).

$$E(r) = 50\% * 10\% + 50\% * 5\% = 7.5\%$$

$$Risk = \sqrt{0.5^2 * 0.16^2 + 0.5^2 * 0} = 8\%$$

$$SR = \frac{7.5-5}{8} = 0.3125$$

Investment B:
Invest 80% in investment B and 20% in the risk free asset:

$$E(r) = 80\% * 8\% + 20\% * 5\% = 7.4\%$$

$$Risk = \sqrt{0.8^2 * 0.1^2 + 0.2^2 * 0} = 8\%$$

$$SR = \frac{7.4-5}{8} = 0.3$$

Investment A, with the higher Sharpe Ratio, should be chosen. For our given level of risk, we can achieve a higher level of return. Notice that each portfolio maintains the Sharpe ratio of the investment when combined with the risk free asset.

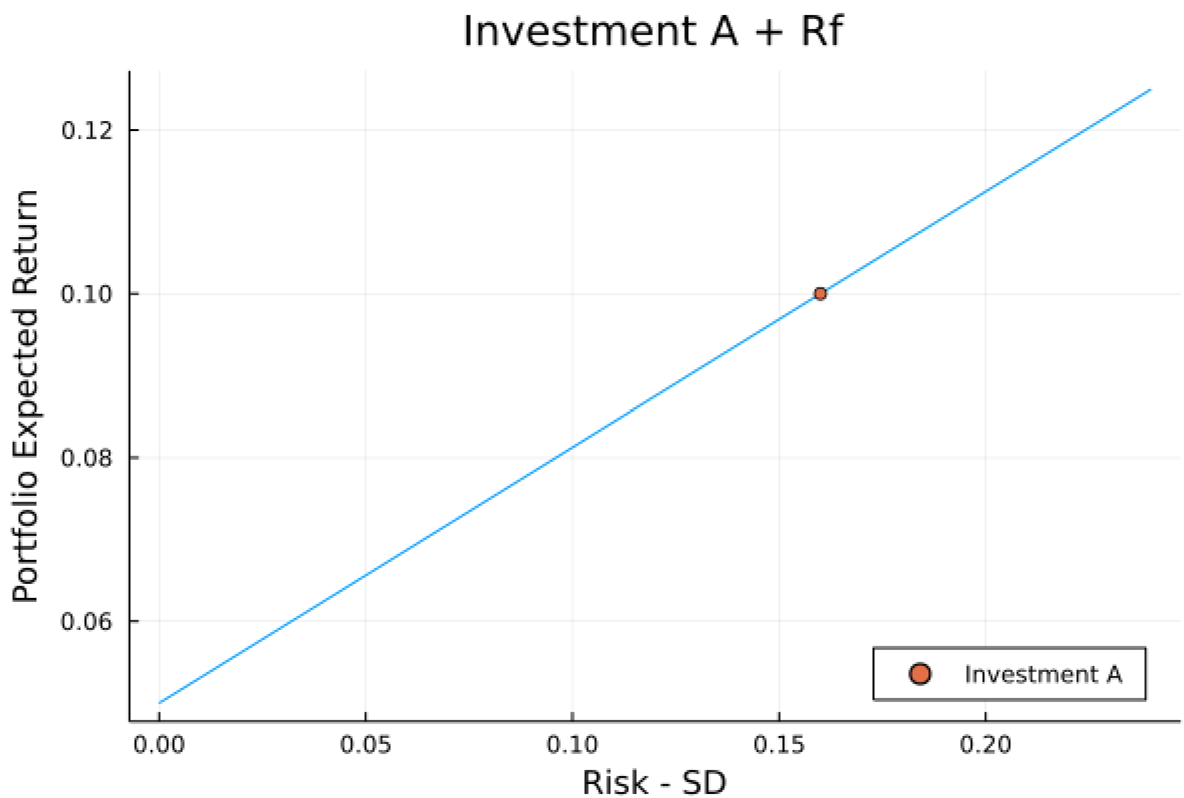
Further, if we wanted to achieve a risk higher than available in investment A, we could always borrow at the risk free rate and use that money to buy additional A.

Borrow 50%

$E(r) = 150\% * 10\% +- 50\% * 5\% = 12.5\%$

$Risk = \sqrt{1.5^2 * 0.16^2 + (-0.5)^2 * 0} = 24\%$

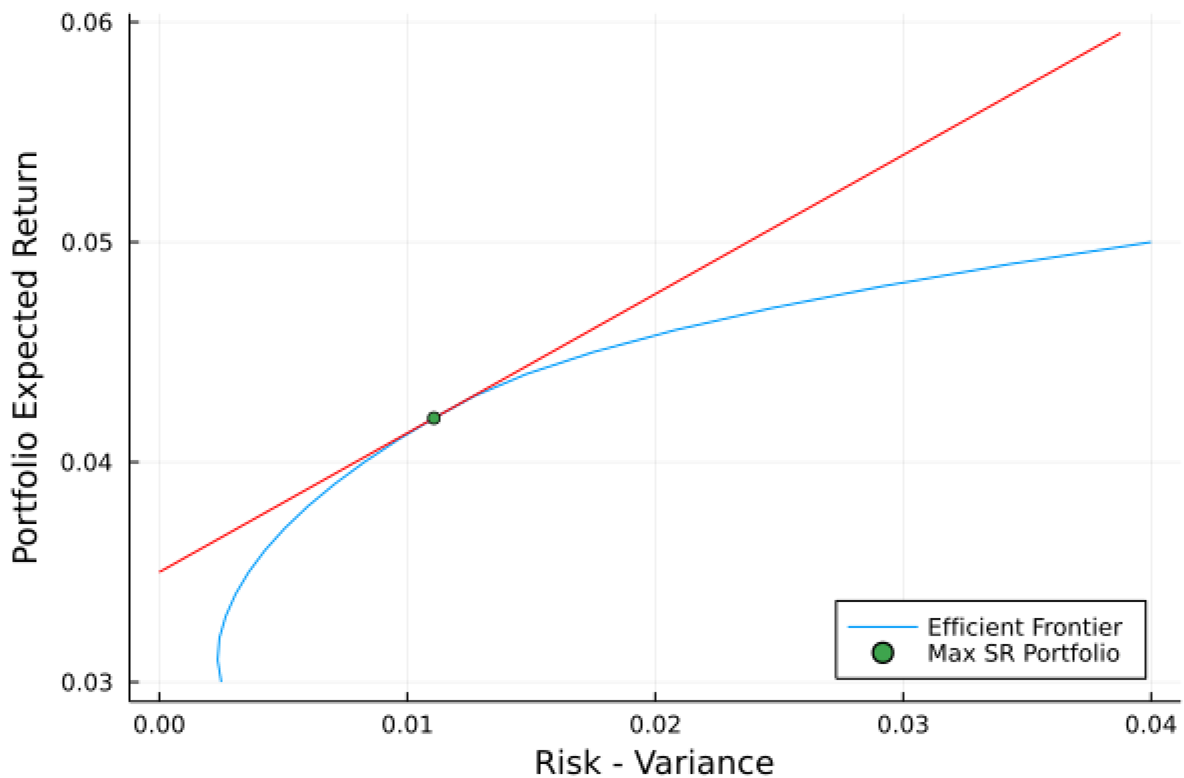
$SR = \frac{12.5-5}{24} = 0.3125$



The intercept of this line is the risk free rate ($\mu = 0.5$, $\sigma = 0$), and the slope is the Sharpe Ratio.

Going back to the efficient frontier from before this begs the question, what is the portfolio with the maximum Sharpe Ratio (assume the risk free is 3.5%)?

Combining the fact that we can create a linear combination of a portfolio with the risk free asset and retain the portfolio's Sharpe Ratio, that means we can construct a portfolio that has a return ≥ the efficient frontier return for all levels of risk.



The red line is called the capital market line (CML). It is tangent to the efficient frontier at the point of the maximum Sharpe Ratio portfolio. This portfolio is sometimes called the super efficient portfolio, or the market portfolio.

Combining CAPM with the Efficient Market Hypothesis (EMH) leads to the conclusion that the market portfolio is the portfolio of all assets weighted by their current market capitalization.

In practice, most investors cannot borrow at the risk free rate. Values on the CML above the market portfolio are unobtainable. Investors either have to borrow at higher rates – lowering the slope the line above the Market Portfolio – or to simply attempt to invest along the efficient frontier.

Under CAPM, rational investors will only invest in the market portfolio and the risk free asset. This presents a single source of systematic risk. The classic CAPM formula for decomposing the returns on a single asset is:

$$r_s - r_{rf} = \alpha + \beta\left(r_{mkt} - r_{rf}\right) + \epsilon_s$$

The excess return on the asset, s, is a linear combination of the excess return on the market portfolio, an alpha value, and an idiosyncratic error term. In the long run, theory would say that $\alpha = 0$.

Arbitrage Pricing Theory

Proposed by Stephen Ross in 1976, Arbitrage Pricing Theory (APT), is a general theory of asset pricing that says financial assets can be modeled as a linear function of various sources of risk and return. These sources, or factors, of risk are systematic and each contribute to an asset's expected return.

$$E\left(r_s - r_{rf}\right) = \alpha_s + \sum_{i=1}^{n} \beta_i E\left(f_i\right)$$

$$E\left(f_i\right) = RP_i$$

Each factor, f, has an expected return, RP, called its Risk Premium. This is the expected return in excess of the risk free rate for taking on the risk of that factor.

CAPM is a APT model with 1 factor – the market risk factor:

$$r_s - r_{rf} = \alpha + \beta\left(r_{mkt} - r_{rf}\right) + \epsilon_s$$

These factors arise for different reasons. As we discussed above, reaching for additional return above the market portfolio along with the inability to borrow at the risk free rate is a leading theory for why certain types of stocks perform better over longer periods which would not be expected given CAPM and EMH. Value stocks and Stocks with low beta values both show long term positive alpha values.

Extensions to CAPM

In 1993 Eugene Fama (Nobel Prize 2013) and Keneth French published a paper which attempted to model known anomalies found when testing CAPM.

Specifically, they added risk factors for Value and Size. It was easy to see in historic returns that stocks with a low Price to Book ratio and stocks with a lower market cap outperformed. Fama and French did not attempt to specify why the anomalies exist, only to model their return.

To do this, they constructed market neutral portfolios for each factor. The value factor would be long value stocks and short expensive stocks. They called this High (Book to Price) minus Low (Book to Price), or HML. The small factor would be long small stocks and short large stocks. This they named small minus big, or SMB.

$$r_s - r_{rf} = \alpha + \beta_{mkt}\left(r_{mkt} - r_{rf}\right) + \beta_{SMB}SMB + \beta_{HML}HML + \epsilon_s$$

In 1997 Mark Carhart published an extension to Fama French adding a momentum factor. Momentum is the tendency of assets that have recently done well to continue doing well. The UMD (up minus down) factor is long high momentum stocks and short low momentum stocks.

$$r_s - r_{rf} = \alpha + \beta_{mkt}\left(r_{mkt} - r_{rf}\right) + \beta_{SMB}SMB + \beta_{HML}HML + \beta_{UMD}UMD + \epsilon_s$$

Over the years, academics have debated these factors, and attempted to find new ones. Most agree that Size is no longer a priced factor. Fama and French argue that their original HML value factor can be decomposed into 2 factors, companies that are profitable RMW and companies that invest in their business CMA.

Some researchers have attempted to revive the value factor by moving away from Book Value. In an age of technology, a lot of a company's value does not accrue to the accounting Book Value. In the last 10+ years value has significantly underperformed leading some researchers to ask if it should be considered a factor.

Quality businesses (RMW and CMA were early attempts at quality) are considered by some to be a priced factor. Like Value, there is no good single definition of Quality.