Credit Risk Analysis Demo

David Donnelly
29 September 2018

Summary

This project is a quick example of a relatively simple credit risk analysis model implemented in C++. This example seeks to model a portfolio of equity investments in firms that are a $[0, 1]$ default option using Merton's model. All the investments are defined as having an initial equity investment and an equity level of default. The risk associated with all the investments will be measured over a varying level of timeframes. Under the new accounting practices defined by CECL, investment risk now has to include both historical and full-term analysis. Previously, banks had only recognized losses when they reached a probable threshold, expanding the timeframe of modeling will capture more systematic risk and make for safer models.

This project has 3 main parts. The first part is a baseline model used to estimate economic parameters of each firm that are necessary in the credit model. The second part is the implementation of the Merton model which uses a mirror of the Black-Scholes option pricing model to estimate the default risk of each investment. Finally, a Monte Carlo simulation is implemented in C++ to run multiple repetitions of the model and generate risk profiles for each firm. None of these steps are up to industry standards (given the short time frame to complete them), but are meant to be a representation of baseline knowledge that can be expanded with better resources and data.

Growth and Volatility Estimation

To start, we'll need some form of analysis to predict the expected future earnings and volatility of each company i. For simplicity, we'll use a model that uses only two predictors: GDP growth in America (Y_t^A) and GDP growth in Europe (Y_t^E) at some time t. We'll define the individual investment i's growth rate ($G_{i,t}$) in time t as:

$$G_{i,t} = w_i^A r_i^A \sqrt{Y_t^A} + w_i^E r_i^E \sqrt{Y_t^E}$$

where w_t^A, w_t^E is the percentage of revenue generated in America and Europe respectively, and r_i^A, r_i^E is a fitted constant for each business to capture how drastically it responds to changes in GDP. This constant would likely be generated by fitting historical data using a method like least squares or a MLE. It's also worth noting that while the decision to choose this formulation is completely arbitrary, a square root function is the right general shape for how this relationship would likely work. The diminishing marginal returns from higher increases and decreases in GDP is likely not an incorrect assumption, but it is just an assumption.

The GDP growth rate for America is modeled using a normal distribution, but the parameters for the distribution depend on the time period t. We want to capture that GDP is likely to be sticky and follow trends through time. In the normal business cycle, GDP will rise and fall over many years, and each year is not independent of the previous. Therefore we define American GDP growth as the following:

$$Y_t^A \sim \mathcal{N}(Y_{t-1}^A, 1.5)$$

This definition means that after each year, the mean expected GDP growth rate is redefined as the actual growth rate over the previous year. Furthermore, while the mean growth rate is changed per year we're making the assumption that the standard deviation around that mean remains constant at 1.5%. Both of these assumptions are simplifications that would be replaced with a more sophisticated method in a more robust model, but for our purposes it will do.

The GDP growth rate for Europe is also modeled by a random variable, but there is a clear relationship between GDP growth in America and GDP growth in Europe. We could either approach the dependency problem by looking for a joint probability function, or use a scaling random variable. Once again we'll make an assumption (that in a real project would be supported by historical data) that European GDP grows about as much as the US, except with a difference that can be captured by a normally distributed random variable:

$$Y_t^E \sim Y_t^A - \mathcal{N}(.1, .05) \sim \mathcal{N}(Y_{t-1}^A, 1.5) - \mathcal{N}(.1, .05)$$

Now that we have the definitions for parameters, we need to formally state the equation for the random variable $G_{i,t}$ and solve for the mean and standard deviation to be used in the credit model.

$$G_{i,t} \sim \begin{cases} w_i^A r_i^A \sqrt{Y_t^A} + w_i^E r_i^E \sqrt{Y_t^E} & Y_t^A > 0 \\ -w_i^A r_i^A \sqrt{\mid Y_t^A \mid} - w_i^E r_i^E \sqrt{\mid Y_t^E \mid} & Y_t^A < 0 \\ 0 & Y_t^A = 0 \end{cases}$$

Now we must solve the expected value and variance for $G_{i,t}$. We'll do it for the positive case knowing the function is perfectly reflected over the Y axis:

$$\mathbb{E}[G_{i,t}] = \mathbb{E}[w_i^A r_i^A \sqrt{Y_t^A} + w_i^E r_i^E \sqrt{Y_t^E}]$$

While we know GDP is not independant between America and Europe, but from our definition of Y_t^E the random aspects of Y_t^A and Y_t^E are independent of one another and Y_t^E is a function of American GDP:

$$\mathbb{E}[G_{i,t}] = \mathbb{E}[w_i^A r_i^A \sqrt{Y_t^A}] + \mathbb{E}[w_i^E r_i^E \sqrt{Y_t^E}]$$

Using the definition of expected value:

$$\mathbb{E}[G_{i,t}] = w_i^A r_i^A \int_0^\infty \frac{\sqrt{x}}{\sqrt{2\pi}} e^{-\sqrt{x}/2} dx + w_i^E r_i^E \int_0^\infty \frac{\sqrt{x}}{\sqrt{2\pi}} e^{-\sqrt{x}/2} dx$$

In place of solving difficult integrals such as this we can use a trick. For some random variable X, that is plugged into a smooth transformation function $g(X)$ (such as the square root at all points other than 0), the expected value can be approximated using a Taylor Expansion around the mean:

$$g(X) = g(\mu) + g'(\mu)(X - \mu) + \frac{g''(\mu)}{2!}(X - \mu)^2 ...$$

Therefore:

$$\mathbb{E}[(g(X)] = \mathbb{E}[g(\mu)] + g'(\mu)\mathbb{E}[(X - \mu)] + \frac{g''(\mu)}{2!}\mathbb{E}[(X - \mu)^2]$$

Noting that becasue our RV Y_t^A is normal:

$$\mathbb{E}[(Y_t^A - \mu)] = 0, \qquad \mathbb{E}[(Y_t^A - \mu)^2] = \sigma_Y^2$$

Using our Taylor expansion and recalling that $\sigma_Y^2 = 2.25$, we can now see that:

$$\mathbb{E}[Y_t^A] \approx \sqrt{Y_{t-1}^A} - \frac{2.25}{8(Y_{t-1}^A)^{3/2}} \approx \sqrt{Y_{t-1}^A} - \frac{.28125}{\sqrt{(Y_{t-1}^A)^3}}$$

$$\mathbb{E}[G_{i,t}] \approx w_i^A r_i^A \left(\sqrt{Y_{t-1}^A} - \frac{.28125}{\sqrt{(Y_{t-1}^A)^3}}\right) + w_i^E r_i^E \left(\sqrt{Y_{t-1}^A} - \frac{.28125}{\sqrt{(Y_{t-1}^A)^3}} - .1\right)$$

This expression can now be used to calculate the mean growth rate for a company that can be used in a credit model. To calculate the variance, we can use the variance definition using expected value. For the first part of the equation ($\mathbb{E}[X^2]$) we don't need the Taylor expansion because it is just the normal expectance for a normal random variable:

$$Var[G_{i,t}] = \mathbb{E}[G_{i,t}^2] - (\mathbb{E}[G_{i,t}])^2$$

$$Var[G_{i,t}] = (w_i^A r_i^A)^2 Y_{t-1}^A + (w_i^E r_i^E)^2 Y_{t-1}^A - \Big(w_i^A r_i^A \Big(\sqrt{Y_{t-1}^A} - \frac{.28125}{\sqrt{(Y_t^A)^3}}\Big) + w_i^E r_i^E \Big(\sqrt{Y_{t-1}^A} - \frac{.28125}{\sqrt{(Y_t^A)^3}} - .1\Big)\Big)^2$$

Rather than simplifying, this ugly algebra, we just programmed it directly into the model. It's important to note that this is only calculated for the positive side of the piecewise equation. A downside of using the square root function is that it's only defined for $Y_t^A \geq 0$. To get around this, we must handle it using a piecewise nature which is easy to accomplish in code.

This completes the first part of the project.We now have a statistically sound prediction method to use to calculate the expected growth and volatility for each asset. The next step is to describe the Merton model we use to measure the default rate.

Merton Credit Model

The Merton credit model is one of the first structural credit models in modern finance, and a powerfully simple credit risk approximater. To start, we treat each firm as financed exclusively through equity E_t and bonds D_t. This would mean the value of the firm V_t can be defined as follows:

$$V_t = E_t + D_t$$

Following the definition we set for our firms above where default occurs only when $V_t < K$ where K is a preset default amount, this would mean E_t could be defined as:

$$E_t = \max\{V_t - K, 0\}$$

Since this is a mirror of the call price for a European option, Merton showed that the equity share E_t can therefore be priced using the Black-Scholes model:

$$E_t = V_t \mathcal{N}(d_+) - K e^{-r(T-t)} \mathcal{N}(d_-)$$

$$d_\pm = \frac{(E_0/K) + (r \pm \sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

Zooming into this equation, we can locate an important aspect of the firm in the second half of the equation. The negative (or loss) side of the equation is the discounted default price multiplied by a probability of loss. While in options the $\mathcal{N}(d_-)$ is usually the probability of the option not hitting the strike price, in the Merton model this is a direct reflection of the default rate, becasue the only way to lose money on this firm is default. We can define:

$$PD_i = \mathcal{N}\Big(\frac{(E_0/K) + (r - \sigma^2)(T-t)}{\sigma\sqrt{T-t}}\Big)$$

Following this estimation technique for default probability, we see that we only need 4 metrics to measure the default risk for any of our firms: E_0, K which would surely be defined for each one, and r, σ. In normal Black-Scholes and option pricing, r is the risk free rate, but in our usage r is not risk free, but instead the expected grown of the firm's equity value. That is exactly what we attempted to estimate in the previous section using $G_{i,t}$. Now putting the two models together we can get a final summary of our default prediction:

$$r_{i,t} = \mathbb{E}[G_{i,t}] \qquad \sigma_{i,t} = Var[G_{i,t}]$$

$$PD_{i,t} = \mathcal{N}\Big(\frac{(E_0/K) + (r_{i,t} - \sigma_{i,t}^2)(T-t)}{\sigma_{i,t}\sqrt{T-t}}\Big)$$

Solution using Monte Carlo

Solving for $PD_{i,t}$ is a difficult task even at this point due to the lack of independence between random variables. Becasue of our assumptions used about the changing economies that define the GDPs and growth rates of each firm, our calculations are clearly not independent across time. Furthermore, to calculate average risk across a portfolio, none of the firms are going to be independent from one another. To address this, we can simply use a Monte Carlo simulation that will run many independent repetitions of a simulated economy, constantly tracking default risk. This is the brunt of the code work, and for further details on implementation see the file "Modeling.cpp."

The results of the Monte Carlo are in the file "Risk.txt". We ran several 10,000 repetition simulations over varying levels of time to see the effect that lifetime risk analysis has on our assets. Sure enough, the longer we measured out our investments the lower the expected default rate got, which was the point of CECL. While a simple model, several different important skills and techniques have been demonstrated in designing and implementing credit modeling in C++, and our results match our intuition about the effect that new accounting regulation has on credit risk analysis.