Regulatory Capital Alpha Factor

Gabriel Magill

2020-06-06

1 Theorems and Definitions

In the following, we assume all the conditions needed for these statements to be true are satisfied.

Random Variable	X
Joint Probability	$P(X,Y)$ $= P(X \cap Y)$ $= P(X\|Y)P(Y)$ $= P(Y\|X)P(X)$
X independent of Y	$\Leftrightarrow P(X \cap Y) = P(X)P(Y)$ $\Leftrightarrow P(X\|Y) = P(X)$
Expectation Value	$\mathbb{E}[X] = \sum_i x_i P(X = x_i) \equiv \sum_i x_i P(x_i)$
Multivariate Expectation Value	$\mathbb{E}[XY] = \sum_{i,j} x_i y_i P(x_i, y_j)$
Conditional Expectation	$\mathbb{E}[X\|Y = y] = \sum_i x_i P(x_i\|y)$
Probability Chain Rule	$P(X \cap (YZ)) = P(X\|YZ)P(YZ)$
Completeness	$\sum_i P(X = x_i) = 1$
Law of total probability	$P(Y) = \sum_i P(Y, x_i)$

2 Motivation

In Basel 2, risk weighted assets (RWA) is the typical amount of the money a counterparty may owe you at some point in the future, accounting for the probability that the counterparty defaults. Under the advanced internal ratings based (AIRB) method, it's calculated as follows:

$$\text{RWA} = \text{EAD} \times (\text{WCDR-PD}) \times \text{LGD} \times \text{MA} \tag{1}$$

Alpha shows up in both standardized and internal OTC EAD formulas:

$$EAD_{\text{SACCR}} = \alpha \cdot (RC + PFE) \tag{2}$$

$$\begin{aligned} EAD_{\text{IMM}} &= \alpha \cdot \text{Effective EPE} \\ \text{Effective EPE} &= \sum_{k=1}^{\text{min(1year, maturity)}} \text{Effective EE}_{t_k} \times \Delta t_k \\ \text{Effective EE}_{t_k} &= \max(\text{Effective EE}_{t_{k-1}}, \text{EE}_{t_k}) \end{aligned} \tag{3}$$

In the IMM and SACCR EAD, α is typically set to 1.4. However, what does α represent, and why do we need it? The following notes attempt to derive RWA capital (denoted below as expectation value of Loss), and outlines the assumptions which give rise to α.

3 Alpha Derivation (1 counterparty)

Let T be some future time. Let $D = \{0, 1\}$ indicate whether the counterparty has defaulted (1) or not (0) at time T. Let E be the exposure at time T. Exposure measures how much money the counterparty owes us, and is defined as $E = \text{Max}(M, 0)$. Here M is the value of the portfolio where a positive number means they owe us and a negative number means we owe them. Note that E and D at future time T are both random variables governed by some probability distribution function. Putting all this together, the amount of money we lose at future time T is:

$$L = D \cdot E \tag{4}$$

In the equation above, we assumed that the loss given default (LGD) is 1, meaning we lose 100% of what we are owed.

Let's now calculate the expectation value of L:

$$\begin{aligned}\mathbb{E}[L] &= \mathbb{E}[D \cdot E] \\ &= \sum_{i,k} e_i d_k P(d_k, e_i) \\ &= \sum_{i,k} e_i d_k P(d_k|e_i) P(e_i)\end{aligned} \tag{5}$$

Let's *assume* that the dependence of defaults on exposure can be captured by a constant fudge factor α. This can be loosely motivated by using Bayes' theorem:

$$\begin{aligned}P(d|e) &= \frac{P(e|d)}{P(e)} P(d) \\ &\approx \alpha P(d)\end{aligned} \tag{6}$$

Saying that α is constant in the equation above violates conservation of probability, but let's call this a minor inconvenience and press onwards. Eq. 5 can then be approximated as

$$\begin{aligned}\mathbb{E}[L] &\approx \sum_{i,k} e_i d_k \alpha P(d_k) P(e_i) \\ &= \left[\sum_i \alpha e_i P(e_i)\right] \left[\sum_k d_k P(d_k)\right] \\ &\equiv \mathrm{EAD} \cdot \mathrm{PD}\end{aligned} \tag{7}$$

Since Eq. 5 approximately equals Eq. 7, we can express α as:

$$\alpha = \frac{\sum_{i,k} e_i d_k P(d_k|e_i) P(e_i)}{\left[\sum_i e_i P(e_i)\right] \left[\sum_i d_k P(d_k)\right]} \tag{8}$$

Therefore, α accounts for the correlation between default probabilities and exposures in the regulatory capital calculation.

4 Alpha Derivation (N counterparties)

Let's assume $N = 2$. The general N case will trivially follow based on the assumptions we will make.

$$
\begin{aligned}
\mathbb{E}[L] &= \mathbb{E}[(D\cdot E)_{\text{cpty A}} + (D\cdot E)'_{\text{cpty B}}] \\
&= \sum_{i,j,k,l} (e_i d_k + e'_j d'_l) P(d_k, d'_l, e_i, e'_j) \\
&= \sum_{i,j,k,l} (e_i d_k + e'_j d'_l) P(d_k, d'_l | e_i, e'_j) P(e_i, e'_j) \\
&= \sum_{i,j,k,l} e_i d_k P(d_k, d'_l | e_i, e'_j) P(e_i, e'_j) + \sum_{i,j,k,l} e'_j d'_l P(d_k, d'_l | e_i, e'_j) P(e_i, e'_j)
\end{aligned} \tag{9}
$$

Let's now make a number of assumptions:

1. Exposures of A and B are independent: $P(e, e') = P(e)P(e')$
2. Defaults of A and B are independent conditional on exposures: $P(d, d'|e, e') = P(d|e, e')P(d'|e, e')$
3. α_A correlates defaults of A on exposures of A and B: $P(d|e, e') \approx \alpha_A P(d)$
4. α is universal: $\alpha_A = \alpha_B = \alpha$

Using all of these strong assumptions, Eq. 9 becomes:

$$
\begin{aligned}
\mathbb{E}[L] &\approx \sum_{i,j,k,l} e_i d_k P(d_k|e_i e'_j) P(d'_l|e_i e'_j) P(e_i) P(e'_j) + \sum_{i,j,k,l} e'_j d'_l P(d_k|e_i e'_j) P(d'_l|e_i e'_j) P(e_i) P(e'_j) \\
&= \sum_{i,j,k,l} e_i d_k \alpha P(d_k) P(d'_l|e_i e'_j) P(e_i) P(e'_j) + \sum_{i,j,k,l} e'_j d'_l P(d_k|e_i e'_j) \alpha P(d'_l) P(e_i) P(e'_j) \\
&= \sum_{i,k} \alpha e_i P(e_i) d_k P(d_k) \sum_{j,l} P(d'_l|e_i e'_j) P(e'_j) + \sum_{j,l} \alpha e'_j P(e'_j) d'_l P(d'_l) \sum_{i,k} P(d_k|e_i e'_j) P(e_i) \\
&= \sum_{i,k} \alpha e_i P(e_i) d_k P(d_k) + \sum_{j,l} \alpha e'_j P(e'_j) d'_l P(d'_l) \\
&= (EAD \cdot PD)_{\text{cpty A}} + (EAD \cdot PD)'_{\text{cpty B}}
\end{aligned} \tag{10}
$$

In the fourth line above, we have used completeness property of probabilities. Since Eq. 9 approximately equals Eq. 10, we can express α in the general case as:

$$
\alpha = \frac{\sum_{i,j,k,l} (e_i d_k + e'_j d'_l) P(d_k, d'_l, e_i, e'_j)}{\sum_{i,k} e_i P(e_i) d_k P(d_k) + \sum_{j,l} e'_j P(e'_j) d'_l P(d'_l)} \tag{11}
$$

The extension to N counterparties follows naturally.

5 Correlating Exposures and Defaults

We've derived α in terms of a joint probability distribution of exposures and defaults. In practice, how do we correlate these two observables? To do so, we need to introduce the framework of copulas, Rieman-Stieltjes integrals and multivariate normal distributions.

We begin by introducing two independent random variables M and Z_i, such that $M \sim N(0,1)$ and $Z \sim N(0,1)$. This means that M and Z_i are sampled from a Gaussian distribution $N(\mu, \sigma^2)$ with mean μ and variance σ^2. Let's construct a new variable X_i:

$$X_i = \sqrt{\rho} M + \sqrt{1-\rho} Z_i \tag{12}$$

This is the usual procedure done in John Hull for example, where X_i is typically thought of as the asset returns for portfolio i, M is a systemic factor affecting all portfolios, and Z_i is an idiosyncratic factor affecting only i. Let's explore the properties of X_i. Given a variable $y \sim N(\mu, \sigma^2)$, one can prove that $ay + b \sim N(a\mu + b, a^2\sigma^2)$. Therefore, $\sqrt{\rho}M$ and $\sqrt{1-\rho}Z_i$ are also normally distributed. It's also true that the sum of two normally distributed random variables results in a normally distributed variable X_i. So what is the mean and standard deviation of X_i?

$$\begin{aligned}
E(X_i) &= E(\sqrt{\rho}M + \sqrt{1-\rho}Z_i) \\
&= \sqrt{\rho}E(M) + \sqrt{1-\rho}E(Z_i) \\
&= 0 + 0 \\
&= 0 \\
Var(X_i) &\equiv cov(X_i, X_i) \\
&= E((X_i - E(X_i))(X_i - E(X_i))) \\
&= E(X_i X_i) \\
&= E(\rho M^2 + 2\sqrt{\rho(1-\rho)}MZ_i + (1-\rho)Z_i^2) \\
&= \rho E(M^2) + 2\sqrt{\rho(1-\rho)}E(M)E(Z_i) + (1-\rho)E(Z_i^2) \\
&= \rho 1 + 0 + (1-\rho)1 \\
&= 1
\end{aligned} \tag{13}$$

In the above, we have used the independence of M and Z_i, as well as the fact that the variance of M is 1. Therefore, $X_i \sim N(0,1)$ as well. Doing the equivalent calculation between X_i and X_j, one obtains $cov(X_i, X_j) = \rho \equiv \Sigma_{ij}$ for $i \neq j$. According to these *notes*, we can convince ourselves that the collection of X_i forms a normal random vector, with Cholesky Decomposition $\Sigma = AA^T$.

Now, let's classify a default as when $X_i < \bar{x}_i$, for some threshold $\bar{x}_i$. Denoting ϕ as the normal CDF, the probability of this happening is:

$$P(X_i < \bar{x}_i) \equiv \int_{-\infty}^{\infty} 1(X_i < \bar{x}_i) d\phi(x) \tag{14}$$

The integral above is the Rieman-Stieltjes (RS) integral. Given a function f and a CDF g, the RS integral

$$\int_a^b f(x)\mathrm{d}g(x) \tag{15}$$

is to be calculated by taking the limit

$$S(P, f, g) = \sum_{j=0}^{n-1} f(c_j)[g(x_{j+1}) - g(x_j)] \tag{16}$$

when the norm of the partition (i.e. length of longest $x_{j+1} - x_j$ subinterval) tends to 0. The $\{x_j\}$ endpoints are the partition, and $c_j \in [x_j, x_{j+1}]$.

Applying this sum to our example for a given partition P which includes $\bar{x}_i$ as one of the endpoints, we have:

$$\begin{aligned} P(X_i < \bar{x}_i) &\equiv \int_{-\infty}^{\infty} 1(X_i < \bar{x}_i) d\phi(x) \\ &= (\phi(x_1) - \phi(x_0)) + (\phi(x_2) - \phi(x_1)) + \cdots + (\phi(\bar{x}_i) - \phi(x_l)) \\ &= \phi(\bar{x}_i) - \phi(x_0) \\ &= \phi(\bar{x}_i) - \phi(-\infty) \\ &= \phi(\bar{x}_i) \\ &\equiv q \end{aligned} \tag{17}$$

This is true when we shrink the partition size to 0. In the above, we used the fact that the terms in the telescoping sum cancel out, that $f(c)$ is 0 when $c > \bar{x}_i$, and that $\phi(-\infty) = 0$ since ϕ is a CDF.

Let's now revisit our derivation of $P(X_i < \bar{x}_i)$, but let's consider X_i to be as in Eq. 12. In particular, let's fix $M = m$. By doing this, $X \sim N(0, 1)$ is no longer true. We have instead $X \sim N(\sqrt{\rho}m, 1 - \rho)$. We now want to calculate:

$$\begin{aligned} P(X_i(Z_i) < \bar{x}_i | M = m) &= P(\sqrt{\rho}m + \sqrt{1-\rho}Z_i < \bar{x}_i | M = m) \\ &= P\left(Z_i < \frac{\bar{x}_i - \sqrt{\rho}m}{\sqrt{1-\rho}} | M = m\right) \\ &= \phi\left(\frac{\bar{x}_i - \sqrt{\rho}m}{\sqrt{1-\rho}}\right) \end{aligned} \tag{18}$$

where in the last line, we have used the fact that $Z_i \sim N(0,1)$ (similar to the derivation in Eq. 17). Now, this probability depends on m, which came from a random variable. Let's now ask what is the value of $P(X_i(Z_i) < \bar{x}_i | M = m)$, averaged over all values of m:

$$\begin{aligned}
E_m(P(X_i(Z_i) < \bar{x}_i | M = m)) &= E_m\left(\phi\left(\frac{\bar{x}_i - \sqrt{\rho}m}{\sqrt{1-\rho}}\right)\right) \\
&= \int_{-\infty}^{\infty} \phi\left(\frac{\bar{x}_i - \sqrt{\rho}m}{\sqrt{1-\rho}}\right) d\phi(m) \\
&= \int_{-\infty}^{\infty} \phi\left(\frac{\bar{x}_i - \sqrt{\rho}m}{\sqrt{1-\rho}}\right) \phi'(m)dm \text{ (since } \phi \text{ is continuous)} \\
&= \int_{-\infty}^{\infty} \phi\left(\frac{\bar{x}_i - \sqrt{\rho}m}{\sqrt{1-\rho}}\right) \frac{e^{-m^2/2}}{\sqrt{2\pi}} dm \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{\bar{x}_i - \sqrt{\rho}m}{\sqrt{1-\rho}}} \frac{e^{-y^2/2}}{\sqrt{2\pi}} \frac{e^{-m^2/2}}{\sqrt{2\pi}} dydm \\
&= \int_{-\infty}^{\bar{x}_i} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz \\
&= \phi(\bar{x}_i)
\end{aligned} \tag{19}$$

We obtain the same thing as in 17! This is somewhat surprising, but not really if you think about it. Note that to show that the $dydm$ integral equals the dz integral, one can take a derivative of each line with respect to $\bar{x}_i$ and use the fundamental theorem of calculus & chain rule. It will then be fairly easy to show equality. This means that the un-differentiated lines are equal up to a constant independent of $\bar{x}_i$. You can then take $\bar{x}_i = \infty$ to show that the constant is 0.

Let's define:

$$\begin{aligned}
g(m) &= P(X_i(Z_i) < \bar{x}_i | M = m)) - E_m(P(X_i(Z_i) < \bar{x}_i | M = m)) \\
&= \phi\left(\frac{\bar{x}_i - \sqrt{\rho}m}{\sqrt{1-\rho}}\right) - \phi(\bar{x}_i)
\end{aligned} \tag{20}$$

By construction and also by the proofs above, $E_m g(m) = 0$. Since $M \sim N(0,1)$, it is true that in 99% of samples, $M > \phi^{-1}(0.01) = \inf(v : \phi(v) \geq 0.01)$. Since ϕ is monotonic, we have that in 99% of cases,

$$\begin{aligned}
g(m) &< g(\phi^{-1}(0.01)) \text{ (99\% CL)} \\
&\equiv WCDR - PD
\end{aligned} \tag{21}$$

where $PD = \phi(\bar{x}_i)$. This is the term that shows up in the RWA capital. Instead of WCDR-PD, one could have defined this measure of default in many other ways, and it's unclear why they chose to subtract PD.

6 Copulas

In practice, we don't measure probability of defaults according to X_i, but rather according to T. So we introduce $F_i(T_i)$, which is the marginal CDF that the counterparty hasn't defaulted by time T_i. So how do we relate $F_i(T_i)$ to all the previous formalism? Copulas (see *notes*).

Copula: A d-dimensional copula, $C : [0,1]^d \to [0,1]$ is a cumulative distribution function (CDF) with uniform marginals.
Sklar's theorem: Consider a d-dimensional CDF, F_n, with marginals $F_1, \cdots, F_n$. Then there exists a copula, C, such that $F_N(t_1, ..., t_n) = C(F_1(t_1), ..., F_n(t_n))$ for all $x_i \in [-\infty, \infty]$ and $i = 1, ..., n$. If F_i is continuous for all i, then C is unique.

So a copula is just a function with n arguments (the marginal CDFs) that completely determine the d-dimentional CDF. This is simple to understand, but surprising that it's true. Let's assume that the full joint multivariate non-normal distribution of t is given by $F_N(t_1, ..., t_n)$ and is continuous, and the full multivariate normal distribution of X_i is given by $\phi_N(X_1, ..., X_n)$ (this *is* continuous). Let's make a bold assumption that the copula of F_N and the copula of ϕ_N are equal. We also introduce $u_i = F_i(t_i)$. We then have:

$$\begin{aligned}
P_F(T_1 < t_1, ..., T_n < t_n) &\equiv F_N(t_1, ..., t_n) \\
&= F_N(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)) \\
&= C_F(u_1, ..., u_n) \\
&= C_\phi(u_1, ..., u_n) \\
&= \phi_N(\phi_1^{-1}(u_1), ..., \phi_n^{-1}(u_n)) \\
&\equiv P_\phi(X_1 < \phi^{-1}(u_1), ..., X_n < \phi_n^{-1}(u_n))
\end{aligned} \tag{22}$$

We don't know what $P_\phi(X_1 < \phi^{-1}(u_1), ..., X_n < \phi_n^{-1}(u_n))$ is in general since X_i are correlated to each other through M. However, conditional on $M = m$, the various $X_i(m)$ are now independent from each other. Therefore

$$P_\phi(X_1 < \phi^{-1}(u_1), ..., X_n < \phi_n^{-1}(u_n) | M = m) = \prod_i P_{\phi_i}(X_i < \phi^{-1}(u_i) | M = m) \tag{23}$$

and finally

$$P_\phi(X_1 < \phi^{-1}(u_1), ..., X_n < \phi_n^{-1}(u_n)) = \int_{-\infty}^{\infty} \prod_i^n P_{\phi_i}(X_i < \phi^{-1}(u_i) | M = m)) d\phi(m) \tag{24}$$

7 Simulations

Going back to α in Eq. 11, we want to properly calculate the numerator $(e_i d_k + e'_j d'_l) P(d_k, d'_l, e_i, e'_j)$. We can approach this as follows under the N-factor model.
Let $X_i = \frac{e_i - \mu(e_i)}{\sigma(e_i)}$ where e_i is our exposure to counterparty i. Let $M \to M_j$ be the returns on stock markets and CDS indices labelled by j, $\sqrt{\rho} \to a_{ij}$ be the correlation between e_i and M_j, $\bar{x}_i = \phi^{-1}(PD) = \phi^{-1}(F_i(t_i))$ be the historical PD of this counterparty (based on credit rating). Assume our simulation is a monte carlo that simulates certain risks factors (FX rates, IR rates, etc) according to a scenario s at future time steps t and re-values the portfolio to obtain $e(s,t)$. If the MC also simulate $M_j = m(s,t)$, then for each s and t, we can multiply $e(s,t)$ by

$$P_\phi(X_1 < \phi^{-1}(u_1), ..., X_n < \phi_n^{-1}(u_n) | M = m(s,t)) \tag{25}$$

We can then proceed to aggregate these probability weighted $e(s,t)$ exposures. To capture the WCDR, the correlations a_i and historical default probabilities PD_i can be calibrated according to historical stress periods.