Factor Models for Default Risk

October 29, 2025

1 Regulatory Background

Per regulatory protocol all sufficiently large banks must calculate Incremental Risk Charge (IRC) as part of their Market Risk RWA. Historically, the IRC model was part of the Basel 2.5 revision published in 2009-2010. The IRC model is complementary to Value-at-Risk (VaR) in the sense that VaR is designed to capture general and specific forms of Market Risk,, but not idiosyncratic issuer risk. The IRC model applies specifically to debt positons in the Trading Book, e.g. corporate and sovereign debt instruments. IRC is usually modeled in a factor based Monte Carlo framework, wherein it is required that banks estimate the 99.9^{th} quantile of the IRC Loss distribution. This metric represents a 1 in 1,000 event wherein the market value of the applicable Trading Book positions are impacted by issuer default or rating migration.

Similar to the IRC model is the Comprehensive Capital Analysis and Review (CCAR) Issuer Default Loss (IDL) model. The IDL model is a key component of the Global Market Shock aspect of the CCAR Program. As with the IRC model the IDL model is designed to capture the risk of issuer default to the Trading Book, but for a broader set of instruments, which include: Corporate Debt, Sovereign Debt, Equities, and Securitized Products. However, the CCAR IDL is usually reported at a less severe confidence level, e.g. 99^{th} quantile from the IDL Loss distribution.

The next round of internalional regulation for the Banking industry coincides the Basel IV Accords, which are likely to be finalized within the next few years (or sooner). As part of the Market Risk component of the Basel IV regulatoions, the Fundamental Review of the Trading Book (FRTB), a Default Risk Charge (DRC) model will be required for all banks of sufficient systemic importance. The DRC model offers yet another example of a factor formulated Monte Carlo based default risk model.

2 Mathematical Prescription

Factor Model Ingredients:

- Factors $F = (F_1, F_2, \cdots, F_m) \in \mathbb{R}^m$ with $\mathbb{E}[F] = 0$ and $Cov(F) = \Sigma \succ 0$
- Factor Loadings $\underline{\beta} = (\beta_1, \beta_2, \cdots, \beta_m) \in \mathbb{R}^m$
- Idiosyncratic Terms $\underline{\epsilon} = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n) \in \mathbb{R}^n$

In the context of Default Risk Modeling, the m factors represent sources of common randomness by which to explain the joint dynamics of a collection of issuer specific credit events, e.g. default events, rating grade migrations, etc. To be realistic it is often assumed that the drivers of credit events are common across all firms, but effect them in bespoke ways, i.e. the loadings on each factor are specific to each firm. However, the behavorial characteristics of each firm give rise to a unique form of idiosyncratic risk that also needs to be taken into account. In this case represented by ϵ_i, which is assumed independent of the common factors.

Bringing these ideas together a firm specific 'latent factor' can be constructed as follows:

$$X_i := \underline{\beta}_i^{'} F + \epsilon_i \; = \; \sum_{j=1}^{m} \beta_{i,j} F_j + \epsilon_i; \quad i = 1, 2, \cdots, n \tag{1}$$

$$(i)\ \mathbb{E}[X_i]=0, \quad (ii)\ Var(X_i)=\underline{\beta}_i^{'}\Sigma\underline{\beta}_i+\sigma^2=\sum_{j,k}\beta_{i,j}\sigma_{j,k}\beta_{i,k}+\sigma^2$$

or in vector form:

$$\underline{X} := B^{'}F+\epsilon; \quad \text{where} \quad B=[\underline{\beta}_1,\underline{\beta}_2,\cdots,\underline{\beta}_n]\in M_{m,n}(\mathbb{R}) \tag{2}$$

$$(i)\ \mathbb{E}[\underline{X}]=0 \qquad (ii)\ Cov(\underline{X})=B^{'}\Sigma B+\sigma^2 I$$

From here we can make some convenient assumptions about the distribution of F and $\underline{\epsilon}$ to make the model tractable. For example,

$$(i)\ F =_D N_m(0,\Sigma) \qquad (ii)\ \underline{\epsilon} =_D N_n(0,\sigma^2 I)$$

In this case it follows that $\underline{X} =_D N_n(0, B^{'}\Sigma B+\sigma^2)$, i.e. the latent factors follow a joint Gaussian distribution.

- $X_i =_D N(0,\underline{\beta}_i^{'}\Sigma\underline{\beta}_i+\sigma^2)$
- $Cov(X_i,X_j)=\underline{\beta}_i^{'}\Sigma\underline{\beta}_j$

Other alternatives include formulating the framework in a multi-variate Student's t distribution, which is desirable if more realistic tail-event clustering[1], is needed in the model. Lastly, there is a rich toolset associated with copulas that can be employed too.

Given a probability of default for the i^{th} issuer, say $p_i \in (0,1)$ the factor model structure can then be used to simulate dependent, or correlated, credit risk events.

$$\mathbb{P}(\text{Default of Issuer i})=\mathbb{P}(\Phi(X_i)\le p_i)=\mathbb{P}(X_i\le\Phi^{-1}(p_i))=p_i \tag{3}$$

$$\mathbb{P}(\text{Default of Issuer i and Default of Issuer j})=\mathbb{P}\left(\Phi(X_i)\le p_i,\Phi(X_j)\le p_j\right) \tag{4}$$

3 Examples

Vasicek (1987) – Single-factor Model

$$X_i := \omega Z+(1-\omega^2)^{1/2}\epsilon_i; \qquad Z,\epsilon_i \overset{\perp}{=}_D N(0,1)\ \ i=1,2,\cdots,n$$

In this case $X_i =_D N(0,1)$ and $Corr(X_i,X_j)=\omega^2$. The model factors can be interpreted as follows: Z represents a global systemic risk factor, e.g. state of global macro-economy, and ϵ_i denotes issuer specific idiosyncratic terms. This model is simple and easy to understand, but arguably only applies when all issuers a nearly homogenous.

Application: Suppose we have a portfolio of homogenous loans each with a Loss in Default (LID) of $\$100K$. We can then model the portfolio default risk as:

$$Loss=\sum_{i=1}^{n}\mathbf{1}_{D_i}\$100K$$

In this case the expected loss is straight forward to compute:

[1]It is well known that the Normal distribution has tail dependency coefficients equal to zero, $\underline{\lambda}_{1,2} := \lim_{p\to 0^+}\mathbb{P}\left(X_1\le Q_{X_1}(p)|X_2\le Q_{X_2}(p)\right)=0$ whenever the correlation coefficient $\rho\in[-1,1)$ i.e. rare events are essentiallly independent.

$$\mathbb{E}[Loss] = \mathbb{E}\left[\sum_{i=1}^{n} \mathbf{1}_{D_i} \$100K\right] = \$100K \cdot \sum_{i=1}^{n} p_i$$

$$\mathbb{E}[Loss^2] = \mathbb{E}\left[|\sum_{i=1}^{n} \mathbf{1}_{D_i} \$100K|^2\right] = (\$100K)^2 \left(\sum_{i=1}^{n} p_i + \sum_{i \neq j} p_{ij}\right)$$

$$p_{ij} := \mathbb{E}[\mathbf{1}_{D_i} \cdot \mathbf{1}_{D_j}] = \mathbb{P}(\text{Default of Issuer i and Default of Issuer j}) = \int_{-\infty}^{\Phi^{-1}(p_i)} \int_{-\infty}^{\Phi^{-1}(p_j)} \phi(x, y; \omega) dx dy$$

However, often times the tail-risk, e.g. Value-at-Risk(0.999) is of interest, but is not obtainable in closed form. Monte Carlo simulation techniques play a fundamental role in the practical application of these models.

Incremental Risk Charge (2012) – Multi-factor Model

It is an industry standard to formulate an IRC model using 3 types of common factors and 1 unique idiosyncratic factor period issuer. The 3 types of common factors usually take the form:

- Common global systemic factor
- Industry factors, e.g. Technology, Healthcare, Agriculture, Manufacturing, etc.
- Region factors, e.g. North America, Europe, Middle East and Africa, Asia, etc.

Note, there are many ways this can be done. For example, the standard Global Industry Classfication System (GICS) or the Bloomberg Industry Classification System (BICS) can be used to create an industry / sector hierarchical structure within the model. Similarly, various region / country taxonomies are avaiable too.

$$X_i := \theta \cdot Z + \gamma_i \cdot I_{Ind(i)} + \beta_i \cdot R_{Reg(i)} + \left(1 - \theta^2 + \gamma_i^2 + \beta_i^2\right)^{1/2} \cdot \epsilon_i \tag{5}$$

Here $Z, I_{Ind(i)}, R_{Reg(i)}, \epsilon_i \overset{\perp}{=}_D N(0, 1)$, where Z represents a global systemic factor, $I_{Ind(\cdot)}$ are industry specific factors, $R_{Reg(\cdot)}$ denote regional specific factors, and ϵ_i are issuer specific idiosyncratic factors.

- $Ind(\cdot)$ denotes an issuer to industry mapping
- $Reg(\cdot)$ represnts an issuer to region mapping

Under this formulation each issuer latent factor is such that $X_i =_D N(0, 1)$. This is an extension of the classic Vasicek model in many ways that offers a more realistic and richer dependency structure.