Adam 2008

1. **The risk modeling**
   Given
   $$y_n = \beta_n^T \mathcal{Z} + \sqrt{1-\beta_n^T\beta_n}\epsilon_n$$
   where $\mathcal{Z} \sim \mathcal{N}(0, I_S)$ and $\epsilon \sim \mathcal{N}(0, I_N)$, the resulting $y_n \sim \mathcal{N}(0,1)$ as shown below
   $$\mathbb{E}\{y_n\} = \sum_i \beta_{n,i}\mathbb{E}\{\mathcal{Z}_i\} + \sqrt{1-\beta_n^T\beta_n}\mathbb{E}\{\epsilon_n\} = 0$$
   $$\begin{aligned} var\{y_n\} &= \mathbb{E}\left\{(y_n - \mathbb{E}\{y_n\})^2\right\} \\ &= \mathbb{E}\left\{(\beta_n^T\mathcal{Z} + \sqrt{1-\beta_n^T\beta_n}\epsilon_n)^2\right\} \\ &= \beta_n^T\beta_n\mathbb{E}\left\{(\mathcal{Z}-0)^2\right\} + (1-\beta_n^T\beta_n)\mathbb{E}\left\{(\epsilon_n-0)^2\right\} \\ &= \beta_n^T\beta_n var\{\mathcal{Z}\} + (1-\beta_n^T\beta_n)var\{\epsilon_n\} \\ &= 1 \end{aligned}$$

2. **Motivation for threshold between different states $H_{c(n)}^c$**
   The motivation is to model discrete probability in the credit state matrix with a continuous distribution such as the gaussian. In this cawe, we want to set $H_{c(n)}^c$ such that
   $$p(H_{c(n)}^{c-1} \leq y_n \leq H_{c(n)}^c) = p_{c(n)}^c$$
   therefore we can write
   $$p(y_n \leq H_{c(n)}^c) = \sum_{\gamma=1}^{c} p_{c(n)}^\gamma \qquad \xrightarrow{y_n\sim\mathcal{N}(0,1)} \qquad \Phi(H_{c(n)}^c) = \sum_{\gamma=1}^{c} p_{c(n)}^\gamma$$

3. **Confidence interval for monte carlo estimation**
   $$p(L_N(\mathcal{Z},\epsilon) \geq l) \in p(L_N(\mathcal{Z},\epsilon) \geq l) \pm CI$$
   Idea is that for the two naive algorithms that are purported to be equivalent, the CI should be approximately the same

GL 2005: Inner Level Optimization

Inner level optimization involves

$$\min_{\theta \geq 0} \{-\theta l + \psi(\theta)\} \qquad \text{where} \qquad \psi(\theta) = \sum_{k=1}^{m} \log(1 + p_k(e^{\theta c_k} - 1))$$

ψ is strictly convex and passes through the origin. The above optimization is equivalent to

$$\theta^* = \begin{cases} \text{solution to } \frac{\partial \psi}{\partial \theta} = l & l > \frac{\partial \psi}{\partial \theta}|_{\theta=0} \\ 0 & \text{otherwise} \end{cases}$$

where

$$\frac{\partial \psi}{\partial \theta}\Big|_{\theta=0} = \sum_{k=1}^{m} \frac{p_k e^{\theta c_k} c_k}{1 + p_k(e^{\theta c_k} - 1)}\Big|_{\theta=0} = \sum_{k=1}^{m} p_k c_k$$

GL 2005: Likelihood Function for Two Level IS

Likelihood for the inner sampling conditioned on Z is given by

$$e^{-\theta_x(Z)L+\psi(\theta_x(Z),Z)} \qquad \text{where} \qquad \psi(\theta) = \sum_{k=1}^{m} \log(1 + p_k(e^{\theta c_k} - 1))$$

The likelihood function for the outer sampling of Z consists of the following change of distribution

$$Z \sim \mathcal{N}(0, I) \qquad \longrightarrow \qquad Z \sim \mathcal{N}(\mu, I)$$

where μ is the twisting parameter for the outer importance sampling such that the resulting shifted normal distribution resembles the zero variance IS distribution, in other words,

$$\mu = \max_z P(L > x|Z = z)e^{\frac{-z^T z}{2}}$$

Then the likelihood for the outer IS is then

$$\begin{aligned}
\frac{\mathcal{N}(0, I)}{\mathcal{N}(\mu, I)} &= \frac{exp(-\frac{1}{2}z^T z)}{exp(-\frac{1}{2}(z-\mu)^T(z-\mu))} \\
&= exp\left(-\frac{1}{2}z^T z + \frac{1}{2}z^T z + z^T\mu - \frac{1}{2}\mu^T\mu\right) \\
&= e^{-\mu^T Z+\mu^T\mu/2}
\end{aligned}$$

Therefore, the estimator for probability of tail event is given by

$$\mathbb{1}_{L>x}e^{-\theta_x(Z)L+\psi(\theta_x(Z),Z)}e^{-\mu^T Z+\mu^T\mu/2}$$

GL 2005: Likelihood Expression for Exponential Twisting

Show

$$\prod_{k=1}^{m} \left(\frac{p_k}{p_{k,\theta}}\right)^{Y_k} \left(\frac{1-p_k}{1-p_{k,\theta}}\right)^{1-Y_k} = e^{-\theta L+\phi(\theta)}$$

where

$$\phi(\theta) = \sum_{k=1}^{m} \log\Big(1 + p_k(e^{\theta c_k} - 1)\Big)$$
$$L = \sum_{k=1}^{m} c_k Y_k$$
$$p_{k,\theta} = \frac{p_k e^{\theta c_k}}{1 + p_k\left(e^{\theta c_k} - 1\right)}$$

Proof.

$$\begin{aligned} lhs &= \exp\left\{\sum_{k=1}^{m} Y_k \log\left(\frac{p_k}{p_{k,\theta}}\right) + (1-Y_k)\log\left(\frac{1-p_k}{1-p_{k,\theta}}\right)\right\} \\ &= \exp\left\{\sum_{k=1}^{m} Y_k \log\Big(1 + p_k(e^{\theta c_k} - 1)\Big) - Y_k\theta c_k + (1-Y_k)\log\Big(1 + p_k\left(e^{\theta c_k} - 1\right)\Big)\right\} = rhs \end{aligned}$$

where

$$1 - p_{k,\theta} = \frac{1-p_k}{1 + p_k\left(e^{\theta c_k} - 1\right)}$$

$\square$

GL 2005: Algorithm

1. Outer Level IS for sysmatic risk Z
   (a) Find shifted parameter μ for outer IS for Z
   (b) Sample $Z \sim \mathcal{N}(\mu, I)$
2. Inner Level IS for each default indicators Y_k
   (a) Calculate conditional default probabilities $p_k(Z)$ for $k = 1, \cdots, m$

$$p_k(Z) = P(Y_k = 1|Z) = p(X_k > x_k|Z) = P(a_k Z + b_k \epsilon_k > \Phi^{-1}(1 - p_k)|Z)$$

   (b) Compute the twisted parameters $\theta_x(Z)$

$$\theta_x(Z) = \begin{cases} \text{solution to } \frac{\partial}{\partial \theta}\psi_m(\theta, Z) = x & \psi'(0) = \mathbb{E}_p\left\{L|Z\right\} = \sum_{k=1}^{m} p_k(Z)c_k < x \\ 0 & otherwise \end{cases}$$

   (c) Compute default indicators (bernoulli) from twisted conditional default probabilities

$$p_{k,\theta_x(Z)}(Z) = \frac{p_k(Z)e^{\theta_x(Z)c_k}}{1 + p_k(Z)(e^{\theta_x(Z)c_k} - 1)} \qquad k = 1, \cdots, m$$

   (d) Compute Loss $L = c_1 Y_1 + \cdots + c_m Y_m$ under twisted distribution
3. Return the estimator of tail proabilities

$$\mathbb{1}_{L>x} e^{\theta_x(Z)L + \psi(\theta_x(Z), Z)} e^{-\mu^T Z + \mu^T \mu/2}$$

Therefore,

$$P(L > x) = \mathbb{E}_{Z \sim \mathcal{N}(\mu, I)\, Y_k \sim p_{k,\theta_x(Z)}} \left\{ \mathbb{1}_{L>x} e^{\theta_x(Z)L + \psi(\theta_x(Z), Z)} e^{-\mu^T Z + \mu^T \mu/2} \right\}$$

GL 2008: Likelihood Function for Two Level IS

Likelihood for the inner sampling conditioned on Z is given by

$$e^{-\theta_x(Z)L+\psi(\theta_x(Z),Z)} \qquad \text{where} \qquad \psi(\theta)=\sum_{k=1}^{m}\log(1+p_k(e^{\theta c_k}-1))$$

The likelihood function for the outer sampling of Z consists of the following change of distribution

$$\mathbf{Z}\sim\mathcal{N}(\mathbf{0},\mathbf{I}) \qquad \longrightarrow \qquad \mathbf{Z}\sim\sum_{i=1}^{K}\lambda_i\mathcal{N}(\boldsymbol{\mu}_i,\mathbf{I})$$

where $K=|\mathcal{S}_q|$ and i denotes some ordering of elements $\mathcal{F}\in\mathcal{S}_q$. $\boldsymbol{\mu}_1,\cdots,\boldsymbol{\mu}_K$ are shifted means for the Gaussian distribution, and $\lambda_1,\cdots,\lambda_K$ where $\sum_i\lambda_i=1$ are the corresponding coefficients for the mixture. Then the likelihood for the outer IS is then

$$\begin{aligned}
\frac{\mathcal{N}(\mathbf{0},\mathbf{I})}{\sum_{i=1}^{K}\lambda_i\mathcal{N}(\boldsymbol{\mu}_i,\mathbf{I})} &= \left(\frac{\sum_{i=1}^{K}\lambda_i\mathcal{N}(\boldsymbol{\mu}_i,\mathbf{I})}{\mathcal{N}(\mathbf{0},\mathbf{I})}\right)^{-1} \\
&= \left(\sum_i\lambda_i\frac{e^{-\frac{1}{2}(\mathbf{Z}-\boldsymbol{\mu}_i)^T(\mathbf{Z}-\boldsymbol{\mu}_i)}}{e^{-\frac{1}{2}\mathbf{Z}^T\mathbf{Z}}}\right)^{-1} \\
&= \left(\sum_i\lambda_i e^{\frac{1}{2}\mathbf{Z}^T\mathbf{Z}-\frac{1}{2}\mathbf{Z}^T\mathbf{Z}+\mathbf{Z}^T\boldsymbol{\mu}_i-\frac{1}{2}\boldsymbol{\mu}_i^T\boldsymbol{\mu}_i}\right)^{-1} \\
&= \left(\sum_{i=1}^{K}\lambda_i e^{\boldsymbol{\mu}_i^T\mathbf{Z}-\frac{1}{2}\boldsymbol{\mu}_i^T\boldsymbol{\mu}_i}\right)^{-1}
\end{aligned}$$

Therefore, the estimator for probability of tail event is given by

$$\mathbb{1}_{L>x}e^{-\theta_x(Z)L+\psi(\theta_x(Z),Z)}\left(\sum_{i=1}^{K}\lambda_i e^{\boldsymbol{\mu}_i^T\mathbf{Z}-\frac{1}{2}\boldsymbol{\mu}_i^T\boldsymbol{\mu}_i}\right)^{-1}$$

Optimization

Want to solve for

$$\max_x \min_y f(x,y)$$

Same as

$$\max_x f(x,\hat{y}(x)) \quad \text{where} \quad \hat{y}(x) = \arg\min_y f(x,y)$$

Simply we write as a function of 1 variable

$$\max_x \hat{f}(x) \quad \text{where} \quad \hat{f}(x) = f(x,\hat{y}(x))$$

Want to compute the 1st order and 2nd order derivatives

$$\hat{f}'(x) = \frac{\partial f}{\partial x} f(x,\hat{y}(x)) + \frac{\partial f}{\partial y} f(x,\hat{y}(x))\hat{y}'(x)$$

Want to compute $\hat{y}'(x)$ first find critical points

$$\frac{\partial f}{\partial y} f(x,y)|_{y=\hat{y}(x)} = 0$$

Solve for the function

$$f_{yx} = \frac{\partial}{\partial x}\frac{\partial}{\partial y} f(x,y)|_{y=\hat{y}(x)}$$

Solve for $\hat{y}'(x)$

$$f_{yx}(x,\hat{y}(x)) + f_{yy}(x,\hat{y}(x))\hat{y}'(x) = 0 \quad \to \quad \hat{y}'(x) = -\frac{f_{yx}(x,\hat{y}(x))}{f_{yy}(x,\hat{y}(x))}$$

Then compute second derivative, i.e. $\hat{f}''(x)$