Monte Carlo Methoden in Kreditrisiko-Management

P Kreditportfolio bestehend aus m Krediten;
Verlustfunktion $L = \sum_{i=1}^{n} L_i$; Die Verluste L_i sind unabhängig bedingt durch einen Vektor Z von ökonomischen Einflussfaktoren.

Gesucht:

$VaR_\alpha(L) = q_\alpha(L)$, $CVaR_\alpha = E(L|L > q_\alpha(L))$, $CVaR_{i,\alpha} = E(L_i|L > q_\alpha(L))$.

Bei Anwendung von Monte Carlo (MC) Simulation tritt das Problem der Simulation von seltenen Ereignissen auf ("rare event simulation")!

ZB. $\alpha = 0,99$. Nur etwas 1% der standard MC Simulationen führt zu einem Verlust L, sodass $L > q_\alpha(L)$.

Standard MC Schätzer:

$$\widehat{CVaR}_\alpha^{(MC)}(L) = \frac{1}{\sum_{i=1}^{n} I_{(q_\alpha,+\infty)}(L_i)} \sum_{i=1}^{n} L_i I_{(q_\alpha,+\infty)}(L_i)$$

wobei L_i der Verlustwert in der i-ten Simulationslauf ist.

$\widehat{CVaR}_\alpha^{(MC)}(L)$ ist sehr instabil, d.h. hat eine sehr hohe Varianz, wenn die Anzahl der Simulationen n nicht sehr sehr groß ist.

Grundlagen von "Importance Sampling"

Sei X eine ZV in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit absolut stetiger Verteilungsfunktion und Dichtefunktion f.

Gesucht: $\theta = E(h(X)) = \int_{-\infty}^{\infty} h(x)f(x)dx$ für eine bekannte Funktion h.

Berechnung der Wahrscheinlichkeit eines Ereignisses A: $h(x) = I_A(x)$.

Berechnung von CVaR: $h(x) = xI_{x>c}(x)$ mit $c = VaR(X)$.

Algorithmus 1 *(Monte Carlo Integration)*

(1) Generiere X_1,X_2,..., X_n unabhängig aus der Dichte f.

(2) Berechne den standard MC Schätzer $\widehat{\theta}_n^{(MC)} = \frac{1}{n}\sum_{i=1}^{n} h(X_i)$.

Aus dem starken Gesetz der großen Zahlen: $\lim_{n\to\infty} \widehat{\theta}_n^{(MC)} = \theta$ gilt fast sicher.

Im Falle von *seltenen Ereignissen* (zB. $h(x) = I_A(x)$, $P(A) << 1$) ist die Konvergenz sehr langsam.

Sei g eine Wahrscheinlichkeitsdichte, sodass $f(x) > 0 \Rightarrow g(x) > 0$.

Wir definieren das *Likelihood Ratio* als: $r(x) := \begin{cases} \frac{f(x)}{g(x)} & g(x) > 0 \\ 0 & g(x) = 0 \end{cases}$

Es gilt:

$$\theta = \int_{-\infty}^{\infty} h(x)r(x)g(x)dx = E_g(h(x)r(x)) \tag{8}$$

Algorithmus 2 *(Importance Sampling)*

(1) Generiere X_1,X_2,..., X_n *unabhängig aus der Dichte* g.

(2) Berechne den IS-Schätzer $\widehat{\theta}_n^{(IS)} = \frac{1}{n}\sum_{i=1}^{n} h(X_i)r(X_i)$.

g heißt "Importance Sampling"-Dichte.

Ziel: Auswahl einer "Importance Sampling"-Dichte, sodass die Variance des IS-Schätzers wesentlich kleiner als die Varianz des standard MC-Schätzers ist.

$$var\left(\widehat{\theta}_n^{(IS)}\right) = \frac{1}{n^2}(E_g(h^2(X)r^2(X)) - \theta^2)$$

$$var\left(\widehat{\theta}_n^{(MC)}\right) = \frac{1}{n^2}(E(h^2(X)) - \theta^2)$$

Theoretisch kann die Varianz des IS-Schätzers auf 0 reduziert werden!

Annahme $h(x) \geq 0, \forall x$.

Für $g^*(x) = f(x)h(x)/E(h(x))$ gilt: $\widehat{\theta}_1^{(IS)} = h(X_1)r(X_1) = E(h(X))$.

Der IS-Schätzer gibt den richtigen Wert nach einer einzigen Simulation!

Sei $h(x) = I_{\{X \geq c\}}(x)$ wobei $c >> E(X)$ (seltenes Ereignis). Es gilt $E(h^2(X)) = P(X \geq c)$ und

$$E_g(h^2(X)r^2(X)) = \int_{-\infty}^{\infty} h^2(x)r^2(x)g(x)dx = E_g(r^2(X); X \geq c) = \quad (9)$$

$$\int_{-\infty}^{\infty} h^2(x)r(x)f(x)dx = \int_{-\infty}^{\infty} h(x)r(x)f(x)dx = E_f(r(X); X \geq c) \quad (10)$$

Das Ziel ist g so auszuwählen, dass $E_g(h^2(X)r^2(X))$ klein wird, oder sodass $r(x)$ für $x \geq c$ klein und das Ereignis $X \geq c$ unter der Dichte g wahrscheinlicher als unter der Dichte f ist.

Exponential tilting: Bestimmung des IS-Dichte für "light tailed" Variablen

Sei $M_x(t)\colon \mathbb{R} \to \mathbb{R}$ die Momentum-generierende Funktion von X:

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

IS-Dichte: $g_t(x) = \frac{e^{tx} f(x)}{M_X(t)}$

Likelihood Ratio: $r_t(x) = \frac{f(x)}{g_t(x)} = M_X(t) e^{-tx}$.

Sei $\mu_t = E_{g_t}(X) = E(X \exp\{tX\})/M_X(t)$.

Wie kann man ein geeignetes t für ein bestimmtes IS Problem ermitteln?

Z.B. für die Schätzung der Tail-Wahrscheinlichkeit?
Das Ziel ist t so zu wählen, dass $E(r(X); X \geq c) = E(I_{X \geq c} M_X(t) e^{-tX})$ klein wird.

$e^{-tx} \leq e^{-tc}$, für $x \geq c$, $t \geq 0 \Rightarrow E(I_{X \geq c} M_X(t) e^{-tX}) \leq M_X(t) e^{-tc}$.

Wir setzten $t = argmin\{M_X(t) e^{-tc}\colon t \geq 0\}$.

Daraus folgt $t = t(c)$ wobei $t(c)$ die Lösung der Gleichung $\mu_t = c$ ist.

(Eine eindeutige Lösung existiert für alle relevanten Werte von c - ohne Beweis).

Exponential Tilting für die Normalverteilung

Sei $X \sim N(0,1)$ mit Dichtefunktion $\phi(x)$.

$g_t(x) = \frac{e^{tx}\phi(x)}{M_X(t)} = \frac{e^{tx}\phi(x)}{e^{t^2/2}} = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}(x-t)^2\}$ und $\mu_t = \frac{E(X \exp\{tX\})}{M_X(t)} = t$

D.h. unter der Verteilung g_t gilt $X \sim N(t,1)$
Die Gleichung $\mu_t = c$ lautet $t = c$.

IS im Falle von Wahrscheinlichkeitsmaßen

Seien f und g Wahrscheinlichkeitsdichten. Definiere zwei Wahrscheinlichkeitsmasse P und Q:

$$P(A) = \int_{x \in A} f(x)dx \text{ und } Q(A) = \int_{x \in A} g(x)dx$$

Die grundelegende Gleichung der IS (8) lautet dann:

$$\theta = E^P(h(X)) = E^Q(h(X)r(X))$$

Analog: Exponential tilting im Fall von Wahrscheinlichkeitsdichten:

Sei X eine ZV in $(\Omega, \mathcal{F}, P)$ sodass $M_X(t) = E^P(\exp\{tX\}) < \infty$, $\forall t$.

Sei $Q_t(A) := E^P\left(\frac{\exp\{tX\}}{M_X(t)}; A\right)$ ein Wahrscheinlichkeitsmaß in $(\Omega, \mathcal{F})$.

Der IS-Algorithmus bleibt gleich:
Simuliere unabhängige Realisierungen von X_i in $(\Omega, \mathcal{F}, Q_t)$ und setze $\hat{\theta}_n^{(IS)} = (1/n)\sum_{i=1}^n X_i r_t(X_i)$ wobei $r_t(X) = M_X(t) exp\{-tX\}$.

Anwendung von IS auf Bernoulli Mischung Modelle

(siehe Glasserman und Li (2003))

Sei $L = \sum_{i=1}^{m} e_i Y_i$ die Verlustfunktion eines Kreditportfolios.

Y_i sind die Verlustindikatoren mit Default-Wahrscheinlichkeit $\bar{p}_i$ und $e_i = (1 - \lambda_i)L_i$ die positiven deterministischen *Exposures* (λ_i sind recovery rates und L_i sind die Kredithöhen), $i = 1, 2, \ldots, m$.

Sei Z ein Vektor von ökonomischen Einflussfaktoren,
sodass $Y_i|Z$ unabhängig sind und $Y_i|(Z = z) \sim Bernoulli(p_i(z))$.

Ziel: Schätzung von $\theta = P(L \geq c)$ mit Hilfe des IS-Ansatzes, für ein gegebenes c, $c >> E(L)$.

Vereinfachter Fall: Y_i sind unabhängig, $i = 1, 2, \ldots, m$.

Sei $\Omega = \{0, 1\}^m$ der Raum der Zustände vom Zufallsvektor Y.

Das Wahrscheinlichleitsmaß P in Ω:

$$P(\{y\}) = \prod_{i=1}^{m} \bar{p}_i^{y_i} (1 - \bar{p}_i)^{1-y_i}, \ y \in \{0, 1\}^m.$$

Die Momentum-generierende Funktion von L: $M_L(t) = \prod_{i=1}^{m} (e^{te_i}\bar{p}_i + 1 - \bar{p}_i)$.

Das Wahrscheinlichkeitsmaß Q_t:

$$Q_t(\{y\}) = \prod_{i=1}^{n} \left(\frac{\exp\{te_i y_i\}}{\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i} \bar{p}_i^{y_i} (1 - \bar{p}_i)^{1-y_i} \right).$$

Seien $\bar{q}_{t,i}$ neue Default-Wahrscheinlichlichkeiten:

$$\bar{q}_{t,i} := \exp\{te_i\}\bar{p}_i / (\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i).$$

Somit gilt:

$$Q_t(\{y\}) = \prod_{i=1}^{m} \bar{q}_i^{y_i} (1 - \bar{q}_i)^{1-y_i}, \; y \in \{0,1\}^m.$$

D.h. nach der exponential tilting sind die Default-Indikatoren unabhängig mit neuen Default-Wahrscheinlichkeiten $\bar{q}_{t,i}$.

$\lim_{t\to\infty} \bar{q}_{t,i} = 1$ und $\lim_{t\to-\infty} \bar{q}_{t,i} = 0 \Rightarrow$
$E^{Q_t}(L)$ nimmt alle Werte in $(0, \sum_{i=1}^{m} e_i)$ an für $t \in \mathbb{R}$.

Für IS-Anwendungen wähle t, sodass $\sum_{i=1}^{m} e_i \bar{q}_{t,i} = c$.

Allgemeiner Fall: Y_i sind unabhängig bedingt durch Z

1. Schritt: Schätzung der bedingten Überschuss-Wahrscheinlichkeit $\theta(z) := P(L \geq c | Z = z)$ für eine gegebene Realisierung z der ökonomischen Faktoren Z, mit Hilfe des im vereinfachten Fall beschriebenen IS-Ansatzes.

Algorithmus 3 *(IS für die bedingte Verlustverteilung)*

(1) *Für ein gegebenes z berechne die bedingtenn Default-Wahrscheinlichkeiten $p_i(z)$ (wie im einfachen Unabhägigkeitsfall) und löse folgende Gleichung:*

$$\sum_{i=1}^{m} \frac{\exp\{te_i\}p_i(z)}{\exp\{te_i\}p_i(z) + 1 - p_i(z)} = c$$

Die Lösung $t = t(c, z)$ gibt den richtigen tilting-Grad.

(2) *Erzeuge n_1 bedingte Realisierungen des Vektors der Default-Indikatoren $(Y_1, \ldots, Y_m)$. Die einzelnen Indikatoren Y_i, werden unabhängig aus $Bernoulli(q_i)$, $i = 1, 2, \ldots, m$, simuliert, wobei*

$$q_i = \frac{\exp\{t(c, z)e_i\}p_i(z)}{\exp\{t(c, z)e_i\}p_i(z) + 1 - p_i(z)}$$

(3) *Sei $M_L(t, z) := \prod[\exp\{te_i\}p_i(z) + 1 - p_i(z)]$ die bedingte Moment-erzeugende Funktion von L. Seien $L^{(1)}$, $L^{(2)}$,...,$L^{(n_1)}$ die n_1 bedingten Realisierungen von L für die n_1 simulierten Realisierun-*

gen von $Y_1, Y_2, \ldots, Y_m$. *Berechne den* IS*-Schätzer für die Tail-Wahrscheinlichkeit der bedingten Verlustverteilung:*

$$\widehat{\theta}_{n_1}^{(IS)}(z) = M_L(t(c,z),z)\frac{1}{n_1}\sum_{j=1}^{n_1} I_{L^{(j)}\geq c}\,\mathsf{exp}\{-t(c,z)L^{(j)}\}L^{(j)}.$$

2. Schritt: Schätzung der unbedingten Überschuss Wahrscheinlichkeit $\theta = P(L \geq c)$.

Naive Vorgangsweise: Erzeuge mehrere Realisierungen z der Einflussfaktoren Z und berechne $\widehat{\theta}_{n_1}^{(IS)}(z)$ für jede dieser Realisierungen. Der gesuchte Schätzer ist der Durchschnittswert der Schätzer $\widehat{\theta}_{n_1}^{(IS)}(z)$ über alle Realisierungen z.
Das ist nicht die beste Lösung, siehe Glasserman und Li (2003).

Bessere Herangehensweise: IS für die Einflussfaktoren.

Annahme: $Z \sim N_p(0, \Sigma)$ (zB. probit-normal Bernoulli Mischung)

Die IS-Dichte g ist die Dichte von $N_p(\mu, \Sigma)$ für einen "neuen" Erwartungswertvektor $\mu \in \mathbb{R}^p$. Eine gute Wahl von μ sollte zu häufigen Realisierungen z die zu höheren bedingten Default-Wahrscheinlichkeiten $p_i(z)$ führen.

Likelihood Ratio:

$$r_\mu(Z) = \frac{\exp\{-\frac{1}{2}Z^t\Sigma^{-1}Z\}}{\exp\{-\frac{1}{2}(Z-\mu)^t\Sigma^{-1}(Z-\mu)\}} = \exp\{-\mu^t\Sigma^{-1}Z + \frac{1}{2}\mu^t\Sigma^1\mu\}$$

Algorithmus 4 *(vollständige IS für Bernoulli Mischung Modelle mit Gauss'schen Faktoren)*

(1) Erzeuge $z_1, z_2, \ldots, z_n \sim N_p(\mu, \Sigma)$ (n ist die Anzahl der Simulationsrunden)

(2) Für jedes z_i berechne $\widehat{\theta}_{n_1}^{(IS)}(z_i)$ wie in Algorithmus 3.

(3) Berechne den IS-Schätzer für die unbedingte Überschuss-Wahrscheinlichkeit:

$$\widehat{\theta}_n^{(IS)} = \frac{1}{n} \sum_{i=1}^{n} r_\mu(z_i) \widehat{\theta}_{n_1}^{(IS)}(z_i)$$

Die Auswahl von μ

μ soll so gewählt werden, dass die Varianz des Schätzers klein ist.

Idee von Glasserman und Li (2003) (Skizze):

$\widehat{\theta}_{n_1}^{(IS)}(z) \approx P(L \geq c|Z = z) \Rightarrow$
Suche eine gute IS-Dichte für die Funktion $z \mapsto P(L \geq c|Z = z)$.

Ansatz:

a) die optimale IS-Dichte g^* ist proportional zu

$$P(L \geq c|Z = z) \exp\{-\frac{1}{2} z^t \Sigma^{-1} z\}.$$

b) die IS-Dichte hat denselben Modus wie die optimale Dichte g^*.

Das führt zu folgendem Optimierungsproblem:

$$\mu = \operatorname{argmax}_z \left\{ P(L \geq c|Z = z) \exp\{-\frac{1}{2} z^t \Sigma^{-1} z\} \right\}.$$

Exakte Lösung ist schwierig weil $P(L \geq c|Z = z)$ ist i.a. nicht in analytischer Form verfügbar.

Siehe Glasserman und Li (2003) für Lösungsansätze.