Risk Measures for Continuous-Time Markov Chains

Darinka Dentcheva and Andrzej Ruszczyński

Stevens Institute of Technology and Rutgers University

Research supported by NSF award DMS-1311978

ICSP 2016, Búzios, Brazil
June 30, 2016

For a finite state space $\mathcal{X}$, we consider a continuous-time Markov chain $\{X_t\}_{0\leq t\leq T}$ with the transition function

$$Q_{t,r}(y|x) = P(X_r = y \,|\, X_t = x),$$

where $x, y \in \mathcal{X}$ and $0 \leq t < r \leq T$. We assume that the transition rates

$$G_t(y|x) = \lim_{\tau\downarrow 0} \frac{1}{\tau}\big[Q_{t,t+\tau}(y|x) - \delta_x(y)\big], \quad x, y \in \mathcal{X},$$

are uniformly bounded for all $0 \leq t \leq T$. Here,

$$\delta_x(y) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$$

The rates constitute the generator $G_t(\cdot|x) : \mathcal{X} \to \mathcal{M}(\mathcal{X})$, where $\mathcal{M}(\mathcal{X})$ is the set of signed measures on $\mathcal{X}$.

Relevant Probability Spaces

- We denote by $\Xi_{t,r}^{\xi}$, where $0 \leq t < r \leq T$, the space of piecewise-constant, right-continuous functions $x : [t, r] \to \mathcal{X}$ with $x_t = \xi$, equipped with the σ-algebra generated by the finite-dimensional cylinders.
- For every $\xi \in \mathcal{X}$, the transition function Q determines a probability measure $P_{t,r}^{\xi}$ on $\Xi_{t,r}^{\xi}$. We write $\left\{X_{\tau}^{t,\xi}\right\}_{t \leq \tau \leq r}$ for a random element on $\Xi_{t,r}^{\xi}$ distributed according to that measure and $\Xi_{t,r} = \bigcup_{\xi \in \mathcal{X}} \Xi_{t,r}^{\xi}$.
- Let $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ be the filtration generated by the process $\{X_t\}_{0 \leq t \leq T}$.
- The space $\mathcal{Z}_t$ contains all bounded $\mathcal{F}_t$-measurable random variables.

For functions $c \in \mathcal{L}_\infty([0, T] \times \mathcal{X})$ and $f \in \mathcal{L}(\mathcal{X})$, we define the total random cost

$$Z_T(c, f) = \int_0^T c_t(X_t)\, dt + f(X_T)$$

For $0 \le t \le r \le T$ and $\xi_t \in \mathcal{X}$, we define

$$I_{t,r}(c) = \int_t^r c_\tau(X_\tau)\, d\tau,$$
$$I_{t,r}^{\xi_t}(c) = \int_t^r c_\tau(X_\tau^{t,\xi_t})\, d\tau.$$

Here $\mathcal{L}_\infty([t, r] \times \mathcal{X})$ and $\mathcal{L}_\infty(\mathcal{X})$ denote the space of measurable, essentially bounded functions $c : [t, r] \times \mathcal{X} \to \mathbb{R}$ or $f : \mathcal{X} \to \mathbb{R}$ respectively equipped with the respective supremum norm.

Conditional Risk Measures

A mapping $\varrho_{t,T} : \mathcal{Z}_T \to \mathcal{Z}_t$, $t \in [0, T]$, is a conditional risk measure.

- It is monotonic if the inequality $Z_T \leq W_T$ implies $\varrho_{t,T}(Z_T) \leq \varrho_{t,T}(W_T)$;
- It is normalized if $\varrho_{t,T}(0) = 0$;
- It is translation invariant if for all $Z_T \in \mathcal{Z}_T$ and all $Z_t \in \mathcal{Z}_t$, the equality $\varrho_{t,T}(Z_t + Z_T) = Z_t + \varrho_{t,T}(Z_T)$ holds;
- It is convex if for all $Z_T, W_T \in \mathcal{Z}_T$ and all $\alpha \in [0, 1]$ we have $\varrho(\alpha Z_T + (1-\alpha)W_T) \leq \alpha\varrho(Z_T) + (1-\alpha)\varrho(W_T)$;
- It is positively homogeneneous if for all $Z_T \in \mathcal{Z}_T$ and all $\gamma \geq 0$, the equality $\varrho(\gamma Z_T) = \gamma\varrho(Z_T)$ holds;
- It is coherent if it is monotonic, translation invariant, convex, and positively homogeneous;
- It has the local property if for all $Z_T \in \mathcal{Z}_T$ and for any event $A \in \mathcal{F}_t$, the relation $\mathbb{1}_A \varrho_{t,T}(Z_T) = \varrho_{t,T}(\mathbb{1}_A Z_T)$ holds.

Markovian Risk Measures

A dynamic risk measure $\varrho = \{\varrho_{t,T}\}_{t\in[0,T]}$ is a collection of conditional risk measures $\varrho_{t,T} : \mathcal{Z}_T \to \mathcal{Z}_t$.

Time Consistency

A dynamic risk measure ϱ is time consistent, if for all $0 \le t \le r \le T$ and all $Z_T \in \mathcal{Z}_T$, the following holds

$$\varrho_{t,T}(Z_T) = \varrho_{t,T}\big(\varrho_{r,T}(Z_T)\big).$$

Definition

A dynamic risk measure ϱ is Markovian, if for all $0 \le t \le T$, all $\xi_{[0,t]}, \xi'_{[0,t]} \in \Xi_{[0,t]}$, the equality $\xi_t = \xi'_t$ implies that, for all bounded measurable functions $c : [t, T] \times \mathcal{X} \to \mathbb{R}$ and $f : \mathcal{X} \to \mathbb{R}$, it holds

$$\varrho_{t,T}\big(I^{\xi_t}_{t,T}(c) + f(X^{t,\xi_t}_T)\big)(\xi_{[0,t]}) = \varrho_{t,T}\big(I^{\xi'_t}_{t,T}(c) + f(X^{t,\xi'_t}_T)\big)(\xi'_{[0,t]}).$$

For Markovian risk measures having the local property, we write

$$v_t(\xi_t) = \varrho_{t,T}\big(I^{\xi_t}_{t,T}(c) + f(X^{t,\xi_t}_T)\big)(\xi_t).$$

Stochastic Conditional Time Consistency

$\xi_{[0,t]}$ - history of the process X up to time t.

Definition

A dynamic risk measure $\varrho = \left\{\varrho_{t,T}\right\}_{t\in[0,T]}$ is stochastically conditionally time consistent (J. Fan & A. Ruszczyński) if
$\forall\ 0 \le t \le r \le T$, all $\xi_{[0,t]} \in \Xi_{[0,t]}$, and all $Z_T, W_T \in \mathcal{Z}_T$,

$$\varrho_{r,T}(Z_T)\,|\,\xi_{[0,t]} \preceq_{\text{st}} \varrho_{r,T}(W_T)\,|\,\xi_{[0,t]} \Rightarrow \varrho_{t,T}(Z_T)(\xi_{[0,t]}) \le \varrho_{t,T}(W_T)(\xi_{[0,t]}).$$

It is strongly stochastically conditionally time consistent, if for any $r_1, r_2 \in [t, T]$, the inequality $\varrho_{r_1,T}(Z_T)\,|\,\xi_{[0,t]} \preceq_{\text{st}} \varrho_{r_2,T}(W_T)\,|\,\xi_{[0,t]}$ implies $\varrho_{t,T}(Z_T)(\xi_{[0,t]}) \le \varrho_{t,T}(W_T)(\xi_{[0,t]})$.

Theorem

If a dynamic risk measure $\varrho = \left\{\varrho_{t,T}\right\}_{t\in[0,T]}$ is stochastically conditionally time consistent, normalized, and has the translation property, then it is time consistent and has the local property.

Recall $v_t(\xi_t) = \varrho_{t,T}\big(I_{t,T}^{\xi_t}(c) + f(X_T^{t,\xi_t})\big)(\xi_t)$.

Theorem

If a dynamic risk measure $\big\{\varrho_{t,T}\big\}_{t\in[0,T]}$ is stochastically conditionally time-consistent, translation invariant, and Markovian, then for every $0 \le t \le r \le T$ and every $\xi_t \in \mathcal{X}$ a functional $\varsigma_{t,r}^{\xi_t} : \mathcal{L}_\infty(\Xi_{t,r}^{\xi_t}, P_{t,r}^{\xi_t}) \to \mathbb{R}$ exists such that for every Z_T we have

$$v_t(\xi_t) = \varsigma_{t,r}^{\xi_t}\big(I_{t,r}^{\xi_t}(c) + v_r(X_r^{t,\xi_t})\big) \qquad (\sharp)$$

Moreover, the functional $\varsigma_{t,r}^{\xi_t}(\cdot)$ is law invariant with respect to the probability measure $P_{t,r}^{\xi_t}$. If ϱ is coherent, then $\varsigma_{t,r}^{\xi_t}(\cdot)$ is a coherent measure of risk.

Transition Risk in Short Intervals. State Consistency

Assumption: $\varsigma^{\xi_t}_{t,r}(\cdot)$ is Lipschitz continuous on $\mathcal{L}_p\big(\Xi^{\xi_t}_{t,r}, P^{\xi_t}_{t,r}\big)$, $p \in [1, \infty)$.

Theorem

If a dynamic risk measure $\left\{\varrho_{t,T}\right\}_{t\in[0,T]}$ is strongly stochastically conditionally time-consistent, translation invariant, and Markovian, then for every $t \in [0, T]$ a functional $\sigma_t : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \mathcal{L}(\mathcal{X}) \to \mathbb{R}$ exists such that for every Z_T, for all $\xi_t \in \mathcal{X}$, and all $r \in [t, T]$ we have

$$v_t(\xi_t) = \int_t^r c_s(\xi_t)\, ds + \sigma_t\big(\xi_t, Q_{t,r}(\,\cdot\,|\xi_t), v_r\big) + \Delta^{\xi_t}_{t,r},$$

where $\left|\Delta^{\xi_t}_{t,r}\right| \leq \text{const} \cdot (r-t)^{\frac{p+1}{p}}$

(i) $\sigma_t(\cdot,\cdot,\cdot)$ is law invariant with respect to the second argument;

(ii) If ϱ is coherent, then $\sigma_t(\xi_t,\cdot,\cdot)$ is a coherent measure of risk;

(iii) For all $x \in \mathcal{X}$ and all $v \in \mathcal{L}(\mathcal{X})$, we have $\sigma_t(x, \delta_x, v) = v(x)$, where δ_x is the Dirac measure at x [state consistency].

Dual Representations

Theorem

If $\sigma_t(x,m,\cdot)$ is a coherent measure of risk, then the following dual representation is true:

$$\sigma_t\big(x,m,v\big) = \max_{\mu\in\mathcal{A}_t(x,m)} \sum_{y\in\mathcal{X}} v(y)\mu(y), \quad v\in\mathcal{L}(\mathcal{X}),$$

where $\mathcal{A}_t(x,m)\subset\mathcal{P}(\mathcal{X})$ is nonempty, convex, closed, and bounded.

Average Value at Risk

$$\sigma(x,m,v) = \min_{\eta\in\mathbb{R}} \Big\{\eta + \frac{1}{\alpha(x)} \sum_{y\in\mathcal{X}} m(y)\max(0, v(y)-\eta)\Big\}, \qquad (1)$$

where $\alpha(x)\in[\alpha_{\min},\alpha_{\max}]\subset(0,1)$.

$$\mathcal{A}(x,m) = \left\{\mu\in\mathcal{P}(\mathcal{X}) : \mu(y)\le\frac{m(y)}{\alpha(x)},\ y\in\mathcal{X}\right\}.$$

After substituting $m=\delta_x$, we obtain state consistency.

Mean–semideviation of order $p \geq 1$

The risk-transition mapping for $\kappa(x) \in [0,1]$ is given by

$$\sigma(x, m, v) = \sum_{y \in \mathcal{X}} m(y) v(y) + \\ \kappa(x) \Bigg(\sum_{y \in \mathcal{X}} m(y) \Big(\max \Big(0, v(y) - \sum_{z \in \mathcal{X}} m(z) v(z) \Big) \Big)^p \Bigg)^{1/p} \quad (2)$$

For $\frac{1}{p} + \frac{1}{q} = 1$, $\sigma(x, m, \cdot)$ has the dual representation with

$$\mathcal{A}(x, m) = \Big\{ \mu \in \mathcal{P}(\mathcal{X}) : \ \exists\, \varphi \in \mathcal{L}_\infty(\mathcal{X}) : \|\varphi\|_q \leq \kappa(x), \ \varphi \geq 0, \\ \mu(y) = m(y) \big(1 + \varphi(y) - \sum_{z \in \mathcal{X}} \varphi(z) m(z) \big), \ \forall y \in \mathcal{X} \Big\}.$$

The mapping σ is state-consistent, because for $m = \delta_x$ we have

$$\sigma(x, \delta_x, v) = v(x) + \kappa(x) \Big(\big(\max(0, v(x) - v(x)) \big)^p \Big)^{1/p} = v(x).$$

Risk Multikernels and Associated Multifunctions

The set $\mathcal{Q}$ consists of stochastic kernels $Q : \mathcal{X} \to \mathcal{P}(\mathcal{X})$.

The multifunction $\mathfrak{M}_\sigma : \mathcal{Q} \rightrightarrows \mathcal{Q}$

With a transition risk mapping $\sigma(x, m, v)$, we associate

$$\mathfrak{M}_\sigma(Q) = \big\{ M \in \mathcal{Q} : M(x) \in \mathcal{A}(x, Q(x)),\ \forall\, x \in \mathcal{X} \big\}.$$

with $\mathcal{A}(\cdot,\cdot)$ as the multifunction featuring in the dual representation of σ.

$I \in \mathcal{Q}$ assigns to each $x \in \mathcal{X}$ the Dirac measure δ_x.
For a state-consistent mapping σ, we have $\mathfrak{M}_\sigma(I) = \{I\}$.

Tangent Cones of Multikernels

The space $\mathfrak{S}$ of signed finite kernels $K : \mathcal{X} \to \mathcal{M}(\mathcal{X})$ is equipped with the norm $\|K\| = \sup_{\substack{-1\leq\varphi(\cdot)\leq 1\\ x\in\mathcal{X}}} \sum_{y\in\mathcal{X}} \varphi(y)K(y|x)$.

For a set $\mathcal{B} \subset \mathfrak{S}$ and $K \in \mathfrak{S}$, we define $d(K, \mathcal{B}) = \inf_{M\in\mathcal{B}} \|K - M\|$.

For two sets $\mathcal{S}_1, \mathcal{S}_2 \subset \mathfrak{S}$, we use the Pompeiu-Hausdorff distance:

$$\text{dist}(\mathcal{S}_1, \mathcal{S}_2) = \sup\Big(\sup_{K\in\mathcal{S}_1} d(K, \mathcal{S}_2), \sup_{K\in\mathcal{S}_2} d(K, \mathcal{S}_1)\Big).$$

Definition

The tangent cone to $\mathcal{Q}$ at I is $\mathcal{T}_{\mathcal{Q}}(I) = \limsup_{\tau\downarrow 0} \frac{1}{\tau}(\mathcal{Q} - I)$,

Equivalent representation due to convexity:

$$\mathcal{T}_{\mathcal{Q}}(I) = \Big\{K \in \mathfrak{S} : \lim_{\tau\downarrow 0} d\Big(K, \frac{1}{\tau}(\mathcal{Q} - I)\Big) = 0\Big\},$$

Differentiation of multikernels

Definition

A multifunction $\mathfrak{M}_\sigma$ is semi-differentiable at the point I in the direction $G \in \mathcal{T}_{\mathcal{Q}}(I)$ if a nonempty set $\mathfrak{D}_\sigma(G) \subset \mathfrak{S}$ exists, such that for every sequence $\varepsilon_n \downarrow 0$ and every sequence $G_n \to G$, $G_n \in \mathcal{T}_{\mathcal{Q}}(I)$,

$$\lim_{n\to\infty} \frac{1}{\varepsilon_n}\big[\mathfrak{M}_\sigma(I + \varepsilon_n G_n) - I\big] = \mathfrak{D}_\sigma(G),$$

The set $\mathfrak{D}_\sigma(G)$ is called the semiderivative of $\mathfrak{M}_\sigma$ at I in the direction G.

Properties

If $\mathfrak{M}_\sigma$ is semi-differentiable at I in direction $G \in \mathcal{T}_{\mathcal{Q}}(I)$, then $\mathfrak{D}_\sigma(G)$ is a closed, convex, and bounded subset of $\mathcal{T}_{\mathcal{Q}}(I)$.

Average Value-at-Risk and Semi-Deviations

Theorem (Average Value-at-Risk)

The multifunction $\mathfrak{M}_\sigma$ associated with the transition risk mapping (1) is semidifferentiable in every direction $G \in \mathcal{T}_{\mathcal{Q}}(I)$ with the semiderivative

$$\mathfrak{D}_\sigma(G) = \left\{ D \in \mathcal{T}_{\mathcal{Q}}(I) : 0 \le D(y|x) \le \frac{G(y|x)}{\alpha(x)} \text{ for } y \neq x,\ D(\mathcal{X}|x) = 0 \right\}.$$

Theorem (Semi-Deviations)

The multifunction $\mathfrak{M}_\sigma$ associated with the transition risk mapping (2) is semidifferentiable in every direction $G \in \mathcal{T}_{\mathcal{Q}}(I)$ with semiderivative

$$\begin{aligned} \mathfrak{D}_\sigma(G) = \Big\{ D \in \mathcal{T}_{\mathcal{Q}}(I) :\ & \exists\ (\Phi \in \mathfrak{S})\ 0 \le \Phi(y|x) \le \kappa(x),\ \forall\ x, y \in \mathcal{X}, \\ & D(y|x) = G(y|x)\big[1 + \Phi(y|x) - \Phi(x|x)\big],\ \forall y \neq x, \\ & D(x|x) = G(x|x) - \sum_{z \in \mathcal{X}} G(z|x)\Phi(z|x),\ \forall x \in \mathcal{X} \Big\}. \end{aligned}$$

Not all transition risk mappings are semi-differentiable

The mapping $\mathfrak{M}_\sigma$ associated with the worst case risk measure

$$\sigma(x, m, v) = \max_{y: m(y) > 0} v(y)$$

is not semidifferentiable in any direction $K \in \mathcal{T}_{\mathcal{Q}}(I)$, unless $K = 0$.

The multikernel $\mathcal{M}$ is defined by the sets

$$\mathcal{A}(x, m) = \{\mu \in \mathcal{P}(\mathcal{X}) : \mu \ll m\}.$$

Therefore, for every $\varepsilon > 0$,

$$\mathcal{A}(x, \delta_x + \varepsilon K(x)) = \{\mu \in \mathcal{P}(\mathcal{X}) : \mu(y) = 0, \text{ whenever } K(y|x) = 0, y \neq x\}.$$

This set does not depend on ε and is different from δ_x for all $K \neq 0$.

Multigenerators

Lemma

If the risk multikernel $\mathfrak{M}_{\sigma_t}$ is semi-differentiable at I in direction G_t then the multifunctions $\mathcal{A}_t(x,\cdot)$ are semi-differentiable at δ_x in directions $G_t(x)$ and their semi-derivatives $\mathfrak{G}_t(x) \in \mathcal{M}(\mathcal{X})$ have the form

$$\mathfrak{G}_t(x) = \left\{D(x) : D \in \mathfrak{D}_\sigma(G_t)\right\}, \quad x \in \mathcal{X}.$$

Definition

The semi-derivatives $\mathfrak{G}_t(x)$, $x \in \mathcal{X}$ is called the risk multigenerator associated with the generator G_t and the risk multikernel $\mathfrak{M}_{\sigma_t}$.

The support functions $s_{\mathfrak{G}_t(x)} : \mathcal{L}_\infty(\mathcal{X}) \to \mathbb{R}$ of the risk multigenerators $\mathfrak{G}_t(x)$ is defined as follows $s_{\mathfrak{G}_t(x)}(v) = \sup_{\lambda \in \mathfrak{G}_t(x)} \sum_{y \in \mathcal{X}} \lambda(y) v(y)$.

The support functions $s_{\mathfrak{G}_t(x)}(\cdot)$ are Lipschitz continuous with a universal Lipschitz constant for all $x \in \mathcal{X}$.

Theorem (BDE)

Suppose a dynamic risk measure $\varrho = \{\varrho_{t,T}\}_{t\in[0,T]}$ is strongly stochastically conditionally time-consistent, coherent, and Markovian, and its risk multikernels $\mathfrak{M}_{\sigma_t}$ are semi-differentiable at I in the directions G_t for $t \in [0, T]$ with the risk multigenerators $\mathfrak{G}_t$ being measurable and uniformly bounded for $t \in [0, 1]$ (in the Pompeiu–Hausdorff sense). Then the risk value functions $v_t(x)$ satisfy the following differential equations

$$\frac{dv_t(x)}{dt} = -c_t(x) - s_{\mathfrak{G}_t(x)}(v_t), \quad t \in [0, T], \quad x \in \mathcal{X}, \tag{3}$$

$$v_T(x) = f(x), \quad x \in \mathcal{X}. \tag{4}$$

In the risk-neutral case and with $c \equiv 0$, the system (3)–(4) reduces to the classical backward Kolmogorov equations for $v_t(x) = \mathbb{E}[f(X_T)|X_t = x]$:

$$\frac{dv_t(x)}{dt} = -\sum_{y\in\mathcal{X}} G_t(y|x)v_t(y), \quad t \in [0, T], \quad x \in \mathcal{X}.$$

Set $\varepsilon_N = \frac{T}{N}$ for a natural number $N > 0$ and $t_i = i\varepsilon_N$, $i = 0, \dots, N$

Discrete-time Markov chain

Transition kernels $Q_i^N(y|x) = Q_{t_i,t_{i+1}}(y|x)$;
cost of a state x at time t_i is $\varepsilon_N c_{i\varepsilon_N}(x)$ and final cost $f(x)$.

Given transition risk mappings $\sigma_t : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \mathcal{L}(\mathcal{X}) \to \mathbb{R}$, we define a sequence of functions $v_{t_i}^N : \mathcal{X} \to \mathbb{R}$

$$v_{t_i}^N(x) = \varepsilon_N c_{t_i}(x) + \sigma_{t_i}\big(x, Q_i^N(x), v_{t_{i+1}}^N\big), \quad x \in \mathcal{X}, \quad i = 0, \dots, N-1,$$

with $v_T^N \equiv f$ and $v^N : [0, T] \times \mathcal{X} \to \mathbb{R}$ defined as

$$v_t^N(x) = \frac{t_{i+1} - t}{t_{i+1} - t_i} v_{t_i}^N(x) + \frac{t - t_i}{t_{i+1} - t_i} v_{t_{i+1}}^N(x), \quad t \in [t_i, t_{i+1}), \quad x \in \mathcal{X}.$$

Discrete-Time Approximation continued

Assumption $\flat$: For all $x \in \mathcal{X}$, the functions $t \mapsto c_t(x)$ and $t \mapsto \mathfrak{G}_t(x)$ are uniformly continuous on $[0, T]$.

We consider the space $\mathcal{W}$ of functions $v : [0, T] \times \mathcal{X} \to \mathbb{R}$, which are continuous with respect to the first argument, equipped with the norm $\|v\| = \max_{x\in\mathcal{X}} \max_{0\le t\le T} \left|v_t(x)\right|$.

Theorem

Under the assumptions of Theorem (BDE) and assumption $\flat$, the functions v^N converge in $\mathcal{W}$ to the unique solution of system (3)–(4).