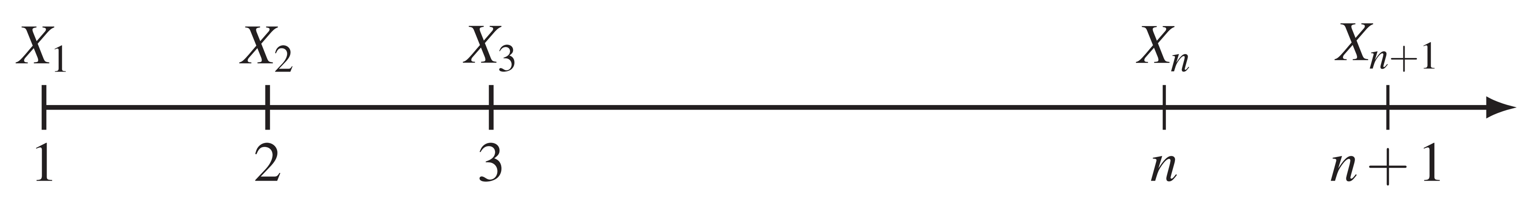
Sum-Product Laws and Efficient Algorithms for Imprecise Markov Chains

Jasper De Bock
Alexander Erreygers
Thomas Krak

GHENT UNIVERSITY
TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

We propose two sum-product laws for imprecise Markov chains, and use these laws to derive two algorithms to efficiently compute lower (and upper) expectations of inferences that have a corresponding sum-product decomposition.

Markov chain P	initial mass function $p_\square$ $p_\square(X_1) = P(X_1)$	transition matrix T_n $T_n(X_n, X_{n+1}) = P(X_{n+1} \mid X_n, X_{1:n-1})$
stochastic process P	$p_\square$	transition matrix $T_{n,X_{1:n-1}}$ $T_{n,X_{1:n-1}}(X_n, X_{n+1}) = P(X_{n+1} \mid X_n, X_{1:n-1})$



consistent **stochastic process** P	$p_\square \in \mathscr{M}_\square$	$T_{n,X_{1:n-1}} \in \mathscr{T}_n$
set **of consistent stochastic processes** $\mathscr{P}$	*set* of initial mass functions $\mathscr{M}_\square$ $\underline{E}_\square\colon \mathbb{R}^{\mathscr{X}} \to \mathbb{R}$ defined for all $f \in \mathbb{R}^{\mathscr{X}}$ by $\underline{E}_\square(f) := \inf_{p_\square \in \mathscr{M}_\square} \sum_{x \in \mathscr{X}} p_\square(x) f(x)$	*set* of transition matrices $\mathscr{T}_n$ $\underline{T}_n\colon \mathbb{R}^{\mathscr{X}} \to \mathbb{R}^{\mathscr{X}}$ defined for all $f \in \mathbb{R}^{\mathscr{X}}$ by $\underline{T}_n f(x) := \inf_{T_n \in \mathscr{T}_n} \sum_{x \in \mathscr{X}} [T_n f](x)$ for all $x \in \mathscr{X}$

Lower (and upper) expectations

For any set $\mathscr{P}$ of consistent stochastic processes, we are interested in lower and upper expectations of the form

$$\underline{E}_{\mathscr{P}}(f(X_{1:n})) := \inf_{P \in \mathscr{P}} E_P(f(X_{1:n}))$$

and

$$\overline{E}_{\mathscr{P}}(f(X_{1:n})) := \sup_{P \in \mathscr{P}} E_P(f(X_{1:n})).$$

Because $\overline{E}_{\mathscr{P}}(f(X_{1:n})) = -\underline{E}_{\mathscr{P}}(-f(X_{1:n}))$, it suffices to study the lower expectation $\underline{E}_{\mathscr{P}}$.

Examples of variables $f(X_{1:n})$ include

★ variables on a single time point

$$g(X_n),$$

★ temporal averages

$$\frac{1}{n}\sum_{k=1}^{n} g(X_k),$$

★ truncated hitting times

$$\min\big(\{k \in \{1,\dots,n\} \colon X_k \in A\} \cup \{n+1\}\big),$$

★ indicators of time-bounded until events,

✪ the number of 'interesting transitions'

$$\big|\{k \in \{1,\dots,n\} \colon (X_{k-1}, X_k) \in A\}\big|,$$

with A a subset of $\mathscr{X}^2$.

A set $\mathscr{P}$ is **compatible with** $\mathscr{M}_\square$ if

$$\underline{E}_{\mathscr{P}}(f(X_1)) = \underline{E}_\square(f) \text{ for all } f \in \mathbb{R}^{\mathscr{X}}.$$

This is the case for $\mathscr{P}^{\mathrm{M}}$ and $\mathscr{P}^{\mathrm{EI}}$.

Imprecise Markov chains

$\mathscr{P}^{\mathrm{M}}$ is the set of *all* consistent Markov chains
$\mathscr{P}^{\mathrm{EI}}$ is the set of *all* consistent stochastic processes

First algorithm

Consider a variable $f(X_{1:n})$ with a ★ first order sum-product decomposition ★

$$f(X_{1:n}) = \sum_{k=1}^{n} g_k(X_k) \prod_{\ell=1}^{k-1} h_\ell(X_\ell),$$

with $h_1, \dots, h_{n-1} \geq 0$.

Let $\underline{\pi}_n := g_n$ and, for all $1 \leq k \leq n-1$,

$$\underline{\pi}_k := g_k + h_k \underline{T}_{k+1} \underline{\pi}_{k+1}.$$

Then in general,

$$\underline{E}_{\mathscr{P}}(f(X_{1:n})) \geq \underline{E}_\square(\underline{\pi}_1).$$

If $\mathscr{P}$ is compatible with $\mathscr{M}_\square$ and satisfies the (first-order) sum-product law, then

$$\underline{E}_{\mathscr{P}}(f(X_{1:n})) = \underline{E}_\square(\underline{\pi}_1).$$

A set $\mathscr{P}$ satisfies the **(first-order) sum-product law** if for all $n \in \mathbb{N}$, $f \in \mathbb{R}^{\mathscr{X}}$ and $g, h \in \mathbb{R}^{\mathscr{X}^n}$ with $h \geq 0$,

$$\underline{E}\big(g(X_{1:n}) + h(X_{1:n}) f(X_{n+1})\big) = \underline{E}\big(g(X_{1:n}) + h(X_{1:n}) \underline{T}_n f(X_n)\big).$$

This is the case for $\mathscr{P}^{\mathrm{M}}$ and $\mathscr{P}^{\mathrm{EI}}$ if for all $n \in \mathbb{N}$, $f \in \mathbb{R}^{\mathscr{X}}$ and $\varepsilon > 0$,

$$(\exists T_n \in \mathscr{T}_n)(\forall x \in \mathscr{X})\ T_n f(x) \leq \underline{T}_n f(x) + \varepsilon.$$

Second algorithm

Consider a variable $f(X_{1:n})$ with a ✪ second order sum-product decomposition ✪

$$f(X_{1:n}) = \sum_{k=2}^{n} g_k(X_{k-1}, X_k) \prod_{\ell=1}^{k-1} h_\ell(X_\ell),$$

with $h_1, \dots, h_{n-1} \geq 0$.

Let $\underline{\pi}_n := 0$ and, for all $1 \leq k \leq n-1$ and $x \in \mathscr{X}$,

$$\underline{\pi}_k(x) := h_k(x)\big[\underline{T}_k\big(g_{k+1}(x,\cdot)\big) + \underline{\pi}_{k+1}\big].$$

Then in general,

$$\underline{E}_{\mathscr{P}}(f(X_{1:n})) \geq \underline{E}_\square(\underline{\pi}_1).$$

If $\mathscr{P}$ is compatible with $\mathscr{M}_\square$ and satisfies the second-order sum-product law, then

$$\underline{E}_{\mathscr{P}}(f(X_{1:n})) = \underline{E}_\square(\underline{\pi}_1).$$

A set $\mathscr{P}$ satisfies the **second-order sum-product law** if for all $n \in \mathbb{N}$, $f \in \mathbb{R}^{\mathscr{X}^2}$ and $g, h \in \mathbb{R}^{\mathscr{X}^n}$ with $h \geq 0$,

$$\underline{E}\big(g(X_{1:n}) + h(X_{1:n}) f(X_n, X_{n+1})\big) = \underline{E}\big(g(X_{1:n}) + h(X_{1:n}) \underline{T}_n f(X_n)\big),$$

with $\underline{T}_n f(x) := [\underline{T}_n(f(x,\cdot))](x)$ for all $x \in \mathscr{X}$.

This is the case for $\mathscr{P}^{\mathrm{M}}$ and $\mathscr{P}^{\mathrm{EI}}$ if for all $n \in \mathbb{N}$, $(f_x)_{x \in \mathscr{X}} \in (\mathbb{R}^{\mathscr{X}})^{|\mathscr{X}|}$ and $\varepsilon > 0$,

$$(\exists T_n \in \mathscr{T}_n)(\forall x \in \mathscr{X})\ T_n f_x(x) \leq \underline{T}_n f_x(x) + \varepsilon.$$