Increasing Convex Monotone Markov Chains: Theory, Algorithm and Applications

Mouad Ben Mamoun[1,2], Ana Bušić[1], Jean-Michel Fourneau[1] and Nihal Pekergin[1,3]

{mobe, abusic, jmf, nih}@prism.uvsq.fr

[1] PR*i*SM, Université de Versailles, France

[2] Université Mohammed V, Rabat, Maroc

[3] Centre Marin Mersenne, Université Paris 1, France

- There are many Markov chains we cannot analyze directly:
  - state space too large
  - only partial information is known (worst case analysis)
- Construction of tractable bounding MC using stochastic comparison
  $\rightarrow$ provides bounds for both transient and steady-state distributions
- Stochastic monotonicity plays an important role in stochastic comparison of Markov chains [Keilson and Kester '77]
  (also [Muller and Stoyan '02])
- Algorithmic stochastic bounding approach:

  Many results for strong stochastic order (st-order) [Abu-Amsha and Vincent '98], [Truffet '01],. . .

  Tutorial: [Fourneau and Pekergin '02]
- Why increasing convex order (icx-order) ?

  Motivation: icx-comparison for random variables is weaker than st-comparison

  Can we obtain better bounds for Markov chains using icx-order?

Outline

Theory

- Stochastic comparison of Markov chains: a brief overview
- Algorithmic characterization of icx-monotonicity on a finite state space

An algorithm for an icx-monotone bounding chain

- Column by column computation: general algorithm
- Some fast solutions for one column
- Small example

Applications

- Worst case arrivals in a Batch/D/1/N queue
- Hierarchical models: bounding PH distributions

Comparison of random variables

$(\mathcal{S}, \leq)$ = a totally ordered space; $\mathcal{F}$ = a family of real functions on $\mathcal{S}$

Def. X and Y random variables on $\mathcal{S}$,

$$X \preceq_{\mathcal{F}} Y \iff E[f(X)] \leq E[f(Y)], \quad \forall f \in \mathcal{F}$$

Strong stochastic order:

$\mathcal{F}_{st}$ = family of all *increasing* real functions on $\mathcal{S}$

Increasing convex order:

$\mathcal{F}_{icx}$ = family of all *increasing and convex* real functions on $\mathcal{S}$

St-comparison is stronger than icx-comparison:

$$\mathcal{F}_{icx} \subset \mathcal{F}_{st} \quad \Longrightarrow \quad (X \preceq_{st} Y \implies X \preceq_{icx} Y)$$

Finite space: $\mathcal{S} = \{1, \ldots, n\}$

Notation: $X \preceq_{\mathcal{F}} Y$ or $x \preceq_{\mathcal{F}} y$, for x and y probability vectors of X and Y

Characterization

- St-comparison: $X \preceq_{st} Y \iff \sum_{k=i}^{n} x_k \leq \sum_{k=i}^{n} y_k, \quad \forall i \in \{1, \ldots, n\}$
- Icx-comparison:
  $X \preceq_{icx} Y \iff \sum_{k=i}^{n}(k-i+1)\, x_k \leq \sum_{k=i}^{n}(k-i+1)\, y_k, \quad \forall i \in \{1, \ldots, n\}$

$$\iff \begin{cases} x_n & \leq & y_n \\ x_{n-1} + 2x_n & \leq & y_{n-1} + 2y_n \\ x_{n-2} + 2x_{n-1} + 3x_n & \leq & y_{n-2} + 2y_{n-1} + 3y_n \\ & \cdots & \\ x_1 + 2x_2 + \ldots + nx_n & \leq & y_1 + 2y_2 + \ldots + ny_n \end{cases}$$

Note: conditions for $i = 1$ are trivial as x and y are probability vectors

$\sum_{k=1}^{n} k\ x_k = \sum_{k=2}^{n}(k-1)\ x_k + 1$

Matrix characterization

$$\boldsymbol{K}_{st} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix}, \quad \boldsymbol{K}_{icx} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 2 & 1 & 0 & \dots & 0 \\ 3 & 2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n & n-1 & n-2 & \dots & 1 \end{pmatrix}$$

$$x \preceq_{st} y \iff x\boldsymbol{K}_{st} \le y\boldsymbol{K}_{st}$$

$$\begin{aligned} x \preceq_{icx} y &\iff x\boldsymbol{K}_{icx} \le y\boldsymbol{K}_{icx} \\ &\iff x\boldsymbol{K}_{icx}(1) \le y\boldsymbol{K}_{icx}(1) \end{aligned}$$

$$\boldsymbol{K}_{icx}(1) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & n-1 & n-2 & \dots & 1 \end{pmatrix}$$

Comparison of Markov chains

Def. $\{X_k\}_{k\geq 0}$ and $\{Y_k\}_{k\geq 0}$ two homogeneous DTMCs:

$$\{X_k\} \preceq_{\mathcal{F}} \{Y_k\} \quad \text{if} \quad X_k \preceq_{\mathcal{F}} Y_k \text{ for all } k \geq 0.$$

Sufficient conditions

1. *Stochastic monotonicity:*

   $\boldsymbol{P}$ transition matrix of $\{X_k\}_{k\geq 0}$

   Def. $\boldsymbol{P}$ is $\mathcal{F}$-monotone if $p \leq_{\mathcal{F}} q \implies p\boldsymbol{P} \leq_{\mathcal{F}} q\boldsymbol{P}$

2. *Comparison of transition matrices:*

   $\boldsymbol{P}$ and $\boldsymbol{Q}$ transition matrices of $\{X_k\}_{k\geq 0}$ and $\{Y_k\}_{k\geq 0}$

   Def. $\boldsymbol{P} \preceq_{\mathcal{F}} \boldsymbol{Q}$ if $P_{i,*} \preceq_{\mathcal{F}} Q_{i,*}, \quad \forall i \in \{1, \ldots, n\}$ $\quad$ ($P_{i,*} = i^{th}$ row of matrix $\boldsymbol{P}$)

Classical theorem for DTMC comparison:

Th. Two homogeneous DTMCs $\{X_k\}_{k\geq 0}$ and $\{Y_k\}_{k\geq 0}$ with transition matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ satisfy $\{X_k\} \preceq_{\mathcal{F}} \{Y_k\}$ if

- $X_0 \preceq_{\mathcal{F}} Y_0$,
- there exists an $\preceq_{\mathcal{F}}$-monotone transition matrix $\boldsymbol{R}$ such that

$$\boldsymbol{P} \preceq_{\mathcal{F}} \boldsymbol{R} \preceq_{\mathcal{F}} \boldsymbol{Q}.$$

If the steady-state distributions $\pi_{\boldsymbol{P}}$ and $\pi_{\boldsymbol{Q}}$ exist, then $\pi_{\boldsymbol{P}} \preceq_{\mathcal{F}} \pi_{\boldsymbol{Q}}$.

Special case: $\boldsymbol{P}$ or $\boldsymbol{Q}$ $\preceq_{\mathcal{F}}$ - monotone.

Def. $\boldsymbol{P}$ is $\mathcal{F}$-monotone if $p \preceq_{icx} q \implies p\boldsymbol{P} \preceq_{icx} q\boldsymbol{P}$

$\rightarrow$ Not suitable for algorithmic approach

Matrix characterization:

$\boldsymbol{P}$ is $\preceq_{icx}$ -monotone $\iff \boldsymbol{K}_{icx}(1)^{-1}\boldsymbol{P}\boldsymbol{K}_{icx}(1) \geq 0$

Note: $\boldsymbol{K}_{icx}$ provides only sufficient conditions

- ${\boldsymbol{K}_{icx}}^{-1}\boldsymbol{P}\boldsymbol{K}_{icx} \geq 0 \Rightarrow \boldsymbol{P}$ is icx-monotone
- There exists an icx-monotone matrix $\boldsymbol{P}$ such that ${\boldsymbol{K}_{icx}}^{-1}\boldsymbol{P}\boldsymbol{K}_{icx} \not\geq 0$

$\boldsymbol{P}$ an arbitrary transition matrix.

Notation:

$$f_{i,j}(\boldsymbol{P}) = \sum_{k=j}^{n}(k-j+1)P_{i,k} = (\boldsymbol{P}\boldsymbol{K}_{icx})_{i,j},\ \forall i \in \{1,\ldots,n\},\ j \in \{2,\ldots,n\}$$

$$f_{*,j}(\boldsymbol{P}) = (f_{i,j}(\boldsymbol{P}))_{i=1}^{n} = (\boldsymbol{P}\boldsymbol{K}_{icx})_{*,j},\ j \in \{2,\ldots,n\}$$

Algorithmic characterization of icx-monotonicity:

$$\boldsymbol{P} \text{ is } \preceq_{icx} \text{-monotone} \iff f_{*,j}(\boldsymbol{P}) \in \mathcal{F}_{icx},\ \forall\ j \in \{2,\ldots,n\}$$

i.e. $\forall\ j \in \{2,\ldots,n\}$, $\begin{cases} f_{1,j}(\boldsymbol{P}) \le f_{2,j}(\boldsymbol{P}) \\ 2f_{i,j}(\boldsymbol{P}) \le f_{i-1,j}(\boldsymbol{P}) + f_{i+1,j}(\boldsymbol{P}),\ \forall i \in \{2,\ldots,n-1\} \end{cases}$

An algorithm for an icx-monotone bounding chain

Problem 1: $\boldsymbol{P}$ an arbitrary transition matrix.
Find an $\preceq_{icx}$-monotone matrix $\boldsymbol{Q}$ such that $\boldsymbol{P} \preceq_{icx} \boldsymbol{Q}$.

Column by column computation (decreasingly)

The constraints for the last column:

$$Q_{i,n} \geq P_{i,n}, \ \forall i$$

$$Q_{*,n} \text{ is increasing and convex}$$

$$Q_{i,n} \leq 1$$

Problem 2: $\mathbf{0} \leq a \leq b$ and b is increasing and convex.
Find an increasing and convex vector $x \in \mathbb{R}^n$ such that $a \leq x \leq b$.

- $Q_{*,n}$ must be a solution of Problem 2 with $a = P_{*,n}$, $b = \mathbf{1}$
- Any solution of Problem 2 can be taken as $Q_{*,n}$

Computation of other columns can be also reduced to Problem 2.

An algorithm for an icx-monotone bounding chain

$\boldsymbol{P}$ an arbitrary transition matrix.

Notation: $s_{i,j}(\boldsymbol{P}) = \sum_{k=j}^{n} P_{i,j} = (\boldsymbol{P}\boldsymbol{K}_{st})_{i,j}, \forall i, j.$

Then $f_{i,j}(\boldsymbol{P}) = f_{i,j+1}(\boldsymbol{P}) + s_{i,j}(\boldsymbol{P}),\ \forall i,\ \forall 2 \leq j < n$

($f_{i,j}(\boldsymbol{P}) = (\boldsymbol{P}\boldsymbol{K}_{icx})_{i,j},\ \forall i,\ \forall 2 \leq j \leq n$)

Algorithm for an icx-monotone bounding chain

- Solve Problem 2 with $a = P_{*,n}$ and $b = 1$. Denote solution by x_n.
  Set $q_{*,n} = s_{*,n}(\boldsymbol{Q}) = f_{*,n}(\boldsymbol{Q}) = x_n$.
- For each $j = n-1$ to 2:
  Solve Problem 2 with vectors a and b:

  $$a_i = \max(f_{i,j}(\boldsymbol{P}), f_{i,j+1}(\boldsymbol{Q}) + s_{i,j+1}(\boldsymbol{Q})),\ b_i = f_{i,j+1}(\boldsymbol{Q}) + 1, \forall i.$$

  Denote the solution by x_j.
  Set $f_{i,j}(\boldsymbol{Q}) = x_j,\ s_{i,j}(\boldsymbol{Q}) = f_{i,j}(\boldsymbol{Q}) - f_{i,j+1}(\boldsymbol{Q}),\ q_{i,j} = s_{i,j}(\boldsymbol{Q}) - s_{i,j+1}(\boldsymbol{Q})$.
- $q_{i,1} = 1 - s_{i,2}(\boldsymbol{Q}), \forall i.$

Properties

For arbitrary transition matrix $\boldsymbol{P}$

- Computation of each column $j,\ 2 \leq n-1$ depends only on 3 vectors:

$$f_{i,j}(\boldsymbol{P}),\ f_{i,j+1}(\boldsymbol{Q}),\ s_{i,j+1}(\boldsymbol{Q})$$

  $\rightarrow$ already known when computing column j.
- There exists always a solution of each of $n-1$ instances of Problem 2.
- If the columns $k \geq j$ are already fixed the conditions on $f_{i,j}(\boldsymbol{Q})$ are both necessary and sufficient.

Th. $\boldsymbol{Q}$ is an $\preceq_{icx}$-monotone matrix such that $\boldsymbol{P} \preceq_{icx} \boldsymbol{Q}$.

An algorithm for an icx-monotone bounding chain

Problem 2: $\mathbf{0} \leq a \leq b$, b is increasing and convex.
Find $x \in \mathbb{R}^n$ such that:

1. $a \leq x \leq b$
2. x is increasing and convex

Existence: trivial solution $x = b$

For column j corresponds to $s_{i,j}(\boldsymbol{Q}) = \sum_{k=j}^{n} q_{i,k} = 1, \forall i \rightarrow$ the worst solution

Optimality: Is there the best solution, i.e. x such that $x \leq y, \forall y$ solution of Problem 2?

Answer: **NO**

Example: $a = (0.1, 0.4, 0.5) \text{ and } b = (1, 1, 1)$

$\bar{y} = (0.1, 0.4, 0.7)$ and $\hat{y} = (0.3, 0.4, 0.5)$ are both solutions of Problem 2

$x \leq \bar{y}$ and $x \leq \hat{y} \Rightarrow x = a$, and a is not convex ($a_3 - a_2 = 0.1 < 0.3 = a_2 - a_1$)

We can consider that a is increasing: if not, take $\bar{a}$ instead of a:

$$\bar{a}_1 = a_1,\ \bar{a}_i = \max(\bar{a}_{i-1}, a_i),\ i > 1$$

Two cases:

1. b is constant vector
2. b is not a constant vector

Linear time complexity solutions of Problem 2

- **Forward computation**:

$$\begin{cases} x_1 = a_1, \\ x_2 = a_2, \\ x_i = \max\{2x_{i-1} - x_{i-2},\ a_i\},\ \forall i \geq 3. \end{cases}$$

  Can yield $x \not\leq b$ even when b is a constant vector

  Example: $a = (0.1, 0.5, 0.5, 0.7)$ and $b = (1, 1, 1, 1) \Rightarrow x = (0.1, 0.5, 0.9, 1.3)$

- **Backward computation**:

$$\begin{cases} x_n = a_n, \\ x_{n-1} = a_{n-1}, \\ x_i = \max\{2x_{i+1} - x_{i+2},\ a_i\},\ \forall i \le n-2. \end{cases}$$

  Property: If b is constant, vector x computed by backward computation is a solution of Problem 2.

  If $\exists\ i$, j such that $i < j$ and $a_i = a_j$, then $a_k = a_j$ for all $k \le j$

  Example: $a = (0, 0, 0.2, 0.2, 0.45, 0.6, 0.6, 0.9)$ and $b = \mathbf{1}$
  $\Rightarrow x = (0.6, 0.6, 0.6, 0.6,\ 0.6,\ 0.6, 0.6, 0.9)$

- **Modified backward computation**: works on segments with the same value and not on individual entries

  Property: If b is constant, vector x computed by modified backward computation is a solution of Problem 2

  Example: $a = (0, 0, 0.2, 0.2, 0.45, 0.6, 0.6, 0.9)$ and $b = \mathbf{1}$
  $\Rightarrow x = (0, 0.075, 0.2, 0.325, 0.45, 0.6, 0.75, 0.9)$

Case when b is not constant

Notation: $\delta = \max_{1 \leq i \leq n}\{b_i - a_i\}, \quad d = (\delta, \dots, \delta)$

If y is a solution of Problem 2 with vectors

$$a + d - b \text{ and } d,$$

then $x = y + b - d$ is a solution of Problem 2 with vectors

$$a \text{ and } b$$

Small example

$$\boldsymbol{P} = \begin{pmatrix} 0.2 & 0 & 0.4 & 0.4 & 0 \\ 0.1 & 0.5 & 0.2 & 0.1 & 0.1 \\ 0.25 & 0.25 & 0 & 0.3 & 0.2 \\ 0.2 & 0.1 & 0 & 0.3 & 0.4 \\ 0.1 & 0 & 0.35 & 0 & 0.55 \end{pmatrix}$$

Exact:	$\pi_{\boldsymbol{P}} = (0.1663, 0.1390, 0.1982, 0.1998, 0.2966)$	$E(\pi_{\boldsymbol{P}}) = 3.3213$
Icx (backward):	$\pi_{\boldsymbol{Q}} = (0.1955, 0.0872, 0.0172, 0.3553, 0.3447)$	$E(\pi_{\boldsymbol{Q}}) = 3.5665$
Icx (mod. backward):	$\pi_{\boldsymbol{Q}'} = (0.1951, 0.1000, 0.0390, 0.3293, 0.3366)$	$E(\pi_{\boldsymbol{Q}'}) = 3.5122$
St:	$\pi_{\boldsymbol{R}} = (0.1111, 0.0544, 0.2302, 0.2594, 0.3449)$	$E(\pi_{\boldsymbol{R}}) = 3.6726.$

For this example:

$$E(\pi_{\boldsymbol{Q}'}) < E(\pi_{\boldsymbol{Q}}) < E(\pi_{\boldsymbol{R}})$$

This is not always the case.

- Worst case analysis

  Example: Worst case arrivals in a Batch/D/1/N queue

  $A = (a_0, \ldots, a_K)$ = distribution of batch arrivals

  Information known: $\alpha = E(A)$ (related to the load - simple to measure)

  We assume: $N >> K$ and $\alpha < 1$

$$
\boldsymbol{P} = \begin{pmatrix}
a_0 & a_1 & \cdots & a_K & 0 & \cdots & 0 \\
a_0 & a_1 & \cdots & a_K & 0 & \cdots & 0 \\
0 & a_0 & a_1 & \cdots & a_K & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \ddots & 0 & a_0 & a_1 & \cdots & a_K \\
\vdots & & & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & 0 & a_0 & \sum_{i=1}^{K} a_i
\end{pmatrix}
$$

  $\mathcal{F}_\alpha$ = the family of all distributions on the space $\{0, \cdots, N\}$ having the mean α

  The icx-worst case distribution: $q = (\frac{N-\alpha}{N}, 0, \ldots, 0, \frac{\alpha}{N})$:

$$
q \in \mathcal{F}_\alpha \quad \text{and} \quad p \preceq_{icx} q, \ \forall p \in \mathcal{F}_\alpha
$$

Construction of an icx-monotone upper bounding matrix:

1. We construct $\boldsymbol{Q}$, an upper icx-bound for $\boldsymbol{P}$ which is not monotone

$$\boldsymbol{Q} = \begin{cases} Q_{0,0} = 1 - \frac{\alpha}{K} & Q_{0,K} = \frac{\alpha}{K} & \\ i = 1, \ldots, N-K+1: & Q_{i,i-1} = (1 - \frac{\alpha}{K}) & Q_{i,i+K-1} = \frac{\alpha}{K} \\ i = N-K+2, \ldots, N-1: & Q_{i,i-1} = (1 - \frac{\alpha}{N-i+1}) & Q_{i,N} = \frac{\alpha}{N-i+1} \\ Q_{N,N-1} = (1-\alpha) & Q_{N,N} = \alpha & \end{cases}$$

2. $t_\delta(\boldsymbol{Q}) = \delta \boldsymbol{Q} + (1-\delta)\boldsymbol{Id}$ (same steady-state distribution)
   $\rightarrow$ allows to move some probability mass to the diagonal elements
3. Then we apply the forward algorithm to the last row of $t_\delta(\boldsymbol{Q})$
4. Finally we change some diagonal and sub-diagonal elements to make the matrix $\preceq_{icx}$-monotone

- Using the fact that the bounding matrix is monotone

  Example: A monotone PH distribution starting in the smallest state is DNBU (Discrete New Better than Used)

  Def. *Discrete New Better than Used (in Expectation) (DNBU(E))*
  X_t an integer valued random variable modeling the residual time of X, given that $X > t$, i.e. $X_t = [X - t | X > t]$. X is said to be:

  - DNBUE if $E(X_t) \leq E(X)$ for all t integer,
  - DNBU if $X_t \preceq_{st} X$ for all t integer.

  DNBU $\Rightarrow$ DNBUE.

  Property: If X is DNBUE of mean m, then $X \preceq_{icx} Y$ where Y is geometric distributed random variable of mean m

- Constructing bounds with specific structure (closed form solutions, easy to analyze numerically, state space reduction ...)

  2-step approach: monotonicity and structure

- It is possible to obtain closer bounds for MCs using icx
- Comparison of st and icx-bounds is still to be done
  $\rightarrow$ detection of classes of problems for which icx-order yields better bounds
- Further optimizations are still possible (Problem 2): we did not fully exploit the fact that we can start by an initial solution $x = b$, iterative methods ?

  $\rightarrow$ compromise between the complexity and the quality of bounds
- Adaptation of algorithms for sparse structures
- Computing icx-bounding chains having special structures