Polynomials of a Stochastic Matrix and Strong Stochastic Bounds

J.M. Fourneau

Laboratoire PRiSM

Université de Versailles (FRANCE)

June 10, 2006

[a] Joint work with Tuğrul Dayar (Bilkent, Turkey), Nihal Pekergin (PRiSM, Versailles, France), and Jean-Marc Vincent (ID, Grenoble, France), supported by a TÜBİTAK-CNRS joint grant, project SurePath from ACI Sécurité, and CNRS-NSF

Motivation

- Solving large Markov chains to check QoS requierements
- Computing bounds rather than exact results
- Strong stochastic bounds on totally ordered state space obtained by some matrix algorithms and numerically solved
- Improving accuracy of the bounds on the steady-state distribution

Strong Stochastic Bounds

- **Definition 1** *Let X and Y be random variables on a* **totally ordered space**, *X is said to be less than Y in the strong stochastic sense ($X \leq_{st} Y$) iff $E[f(X)] \leq E[f(Y)]$ for all nondecreasing functions f whenever the expectations exist.*

- Average number of customers, loss rates are non decreasing functions.

- Discrete Time Markov Chains (continuous time MC can be considered after uniformization)

- **Definition 2** *Let X and Y be random variables on the finite state space $\{1, 2, \ldots, n\}$. Let p and q be probability distribution vectors such that*

$$p_j = Pr(X = j) \quad and \quad q_j = Pr(Y = j) \quad for \quad j = 1, 2, \ldots, n.$$

X is said to be less than Y in the strong stochastic sense, that is, $X \leq_{st} Y$ iff

$$\sum_{j=k}^{n} p_j \leq \sum_{j=k}^{n} q_j \quad for \quad k = 1, 2, \ldots, n.$$

- $(0.1, 0.2, 0.4, 0.3) <_{st} (0.05, 0.25, 0.2, 0.5)$ because

$0.3 \leq 0.5$ *and* $0.4 + 0.3 \leq 0.2 + 0.5$ *and* $0.2 + 0.4 + 0.3 \leq 0.25 + 0.2 + 0.5$

Fundamental Theorem

- $P_{i,*}$ refers to row i of P.
- **Theorem 1** *Let P and Q be stochastic matrices respectively characterizing time-homogeneous MCs $X(t)$ and $Y(t)$. Then* $\{X(t),\ t \in \mathcal{T}\} \leq_{st} \{Y(t),\ t \in \mathcal{T}\}$ *if*
  - $X(0) \leq_{st} Y(0)$,
  - *st-monotonicity of at least one of the matrices holds, that is,*

    $$\textit{either} \quad P_{i,*} \leq_{st} P_{j,*} \quad \textit{or} \quad Q_{i,*} \leq_{st} Q_{j,*} \quad \forall i,j \quad \textit{such that} \quad i \leq j,$$

  - *st-comparability of the matrices holds, that is,* $P_{i,*} \leq_{st} Q_{i,*}\ \ \forall i$.
- we obtain several algorithms based on monotonicity, comparability and some properties to improve but speed of the resolution (lumpability)

Improving the accuracy of the bounds

Possible Solutions

- Reordering the states (Dayar and Pekergin, EUJOR)
- Linear transform (Dayar, Fourneau Pekergin, RAIRO)
- Polynomial transform.

An algebraic approach

- **Definition 3** *Let $\mathcal{D}$ be the set of polynomials $\Phi()$ such that $\Phi(1) = 1$, Φ different of Identity, and all the coefficients of Φ are non negative.*

- **Lemma 1** *$\Phi(P)$ has the same steady-state distribution than P.*

- **Lemma 2** *Let P be an ergodic chain, and let $\Phi()$ be an arbitrary polynomial in $\mathcal{D}$, $\Phi(P)$ is ergodic.*

- $<_{el}$ is the element-wise comparison.

Operators

- **Definition 4** *Let P be an arbitrary stochastic matrix, we define:*
  - r*: the summation operator.*

$$r(P)[i,j] = \sum_{k=j}^{n} P[i,k]$$

  - v *:*

$$v(P)[1,j] = \sum_{k=j}^{n} P[1,k] = r(P)[1,j]$$

And,

$$v(P)[i,j] = max_{m \leq i}\{\sum_{k=j}^{n} P[m,k]\} = max_{m \leq i}\{r(P)[m,j]\}$$

- **Property 1** *The strong stochastic comparison and the monotonicity of stochastic matrices can be defined through operator* r*:*

  a- st-comparability:

  $$P1 <_{st} P2 \quad \Leftrightarrow \quad r(P1) <_{el} r(P2) \; = \; r(P1)[i,j] \leq r(P2)[i,j] \quad \forall i,j$$

  b- $P1$ *is st-monotone*

  $$\Leftrightarrow \quad r(P1)[i,j] \leq r(P1)[i+1,j] \quad 1 \leq i \leq n-1 \;\; \textit{and} \;\; \forall j$$

- **Proposition 1** *The operator defined by Vincent's algorithm is* $r^{-1}v$*.*

Optimality of Vincent's algorithm

- **Lemma 3** *Let* $M = r^{-1}v(P)$ *then* M *is the smallest st-monotone stochastic matrix larger than* P, *i.e. let* Q *be an arbitrary stochastic matrix such that*
  - Q *is st-monotone*
  - $P <_{st} Q$

  thus M *is smaller than* Q, *i.e.* $M <_{st} Q$
- Thus the matrix obtained by Vincent's algorithm is the smallest one (using the st comparison of matrices).
- But the distribution is not the most accurate.

Fundamental properties: stability

- **Lemma 4** *Let* Φ *be an arbitrary polynomial in* $\mathcal{D}$*, let* P *et* Q *be two stochastic matrices in* $\mathcal{A}$*, if* $P <_{st} Q$ *and* Q *st-monotone then*
  1. $\Phi(Q)$ *is st-monotone*
  2. $\Phi(P) <_{st} \Phi(Q)$
- because if Q is st-monotone, then Q^i is also st-monotone for any positive value of i.
- and the convex sum of st-monotone stochastic matrices is st-monotone

Main results

- **Theorem 2** *Let Φ be an arbitrary polynomial in $\mathcal{D}$, we have*

$$r^{-1}v(\Phi(P)) <_{st} \Phi(r^{-1}v(P))$$

- **Corollary 1** *Consider an arbitrary polynomial $\Phi()$ in $\mathcal{D}$ for an arbitrary ergodic Markov chain P, we have :*

$$\pi_P <_{st} \pi_{r^{-1}v(\Phi(P))} <_{st} \pi_{r^{-1}v(P)}$$

- the pre-processing by $\Phi()$ increases the accuracy.

Toy Example

- we study the polynomials $\phi(X) = X/2 + 1/2$ and $\psi(X) = X^2/2 + 1/2$
- $P = \begin{pmatrix} 0.1 & 0.2 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.2 & 0.3 \\ 0.1 & 0.5 & 0.4 & 0 \\ 0.2 & 0.1 & 0.3 & 0.4 \end{pmatrix}$
- $\phi(P) = \begin{pmatrix} 0.55 & 0.1 & 0.2 & 0.15 \\ 0.1 & 0.65 & 0.1 & 0.15 \\ 0.05 & 0.25 & 0.7 & 0 \\ 0.1 & 0.05 & 0.15 & 0.7 \end{pmatrix}$ $\psi(P) = \begin{pmatrix} 0.575 & 0.155 & 0.165 & 0.105 \\ 0.08 & 0.63 & 0.155 & 0.135 \\ 0.075 & 0.185 & 0.65 & 0.09 \\ 0.075 & 0.13 & 0.17 & 0.625 \end{pmatrix}$

  $r^{-1}v(\phi(P)) = \begin{pmatrix} 0.55 & 0.1 & 0.2 & 0.15 \\ 0.1 & 0.55 & 0.2 & 0.15 \\ 0.05 & 0.25 & 0.55 & 0.15 \\ 0.05 & 0.1 & 0.15 & 0.7 \end{pmatrix}$ $r^{-1}v(P) = \begin{pmatrix} 0.1 & 0.2 & 0.4 & 0.3 \\ 0.1 & 0.2 & 0.4 & 0.3 \\ 0.1 & 0.2 & 0.4 & 0.3 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{pmatrix}$

  $r^{-1}v(\psi(P)) = \begin{pmatrix} 0.575 & 0.155 & 0.165 & 0.105 \\ 0.08 & 0.63 & 0.155 & 0.135 \\ 0.075 & 0.185 & 0.605 & 0.135 \\ 0.075 & 0.13 & 0.17 & 0.625 \end{pmatrix}$

Toy Example-continued

- Finally, we compute the steady-state distribution for the bounds and the initial matrix :

$$\pi_P = (0.1530, 0.3025, 0, 3167, 0.2278)$$

$$\pi_{r^{-1}v(P)} = (0.1, 0.2, 0, 3667, 0.3333)$$

$$\pi_{r^{-1}v\phi(P)} = (0.1259, 0.2587, 0, 2821, 0.3333)$$

$$\pi_{r^{-1}v\psi(P)} = (0.1530, 0.2997, 0, 2916, 0.2557)$$

- $\psi()$ gives the better bound.

How to find a good polynomial

- For degree 1 polynomial, it is proved that the bound does not improve anymore when the matrix becomes Raw Diagonal Dominant
- Thus $\phi(X) = X/2 + 1/2$ is the best among the degree 1 polynomials
- Toy example shows that a degree 2 polynomial is even better.
- Question : Can we find the best polynomial for any degree ?
- Question : Tradeoff between complexity and accuracy ?
- Question : Higher the degree, more accurate the bound ?

The set of good polynomials

- **Definition 5** *Let P be a stochastic matrix, and $k \geq 1$, Ψ is a good polynomial with degree k iff $\pi_{r^{-1}v(\Psi(P))} \leq_{st} \pi_{r^{-1}v(\Phi(P))}$ for all polynomial $\Phi()$ with degree $\leq k$.*

- Good polynomials of degree k for a chain yield the most accurate strong stochastic bound computed by Vincent's algorithm on the steady-state distribution of this chain.

- **Theorem 3** *The set of good polynomials of degree n is empty or convex*

Increasing the degree is not always a good idea

- Idea: P^∞ is monotone
- But increasing the degree is not always a good idea
- For instance for antimonotone Markov chain.
- **Definition 6** *A Markov chain with matrix P is antimonotone, iff $\forall$ u and v such that $u <_{st} v$ then $vP <_{st} uP$.*
- **Property 2** *If P is antimonotone then P^2 is monotone*
- **Property 3** *If P is monotone then $\Phi(P)$ is monotone*
- **Property 4** *If P is antimontone, all the polynomials with odd coefficients equal to 0 are good polynomials and they provide a bound equal to the exact solution.*

Conclusion

- Pre-processing of the stochastic matrix to improve accuracy
- Tradeoff between complexity and accuracy
- An algebraic approach but a stochastic version is in the paper.
- A tool available
- Several models (loss rates and queueing delays)