NON-ASYMPTOTIC APPROXIMATION of Markov Chain Functionals in Transient Regimes

Manuel E. Lladser
University of Colorado - Boulder

In collaboration with J. Barrera

AofA 2017

Problem Setting and Statement.

- S: **finite** but possibly **very large state space**
- $X \sim Markov(\mu, p)$ i.e. a first-order homogeneous Markov chain:
  - μ : **initial distribution**
  - p : **probability transition matrix**

Problem.

Approximate the distribution of functionals of the form:

$$F_n := \sum_{t=1}^{n} f(t, X_t)$$

when n is **too large for exact calculations** [1] and **too small to rely on the Normal** [2] or **the Poisson approximation** [3]

[1] Durrett'99, Flajolet-Sedgewick'09

[2] Bender-Kochman'93, Régnier-Szpankowski'98, Nicodème-Salvy-Flajolet'02

[3] Aldous'88, Barbour-Holst-Janson'92, Erhardsson'99, Roquain-Schbath'07

Problem Setting and Statement. (Talk Simplification!)

- S: **finite** but possibly **very large state space**
- $X \sim Markov(\mu, p)$ i.e. a first order homogeneous Markov chain:
  - μ : **initial distribution**
  - p : **probability transition matrix**

Problem.

Approximate the distribution of functionals of the form:

$$T_n := \sum_{t=1}^{n} [\![X_t \in T]\!] \quad (\textbf{sojourn-time} \text{ of a set } T)$$

when n is **too large for exact calculations** [1] and **too small to rely on the Normal** [2] or **the Poisson approximation** [3]

[1] Durrett'99, Flajolet-Sedgewick'09

[2] Bender-Kochman'93, Régnier-Szpankowski'98, Nicodème-Salvy-Flajolet'02

[3] Aldous'88, Barbour-Holst-Janson'92, Erhardsson'99, Roquain-Schbath'07

Moment Generating Function of T_n.

- **Symbolic Method:** z marks # of visits to T
- **Transfer Matrix** [1]**:** $p_z(i,j) := p(i,j) \cdot z^{[\![j\in T]\!]}$
- **Moment Generating Function** [2]**:** $\mathbb{E}\left(z^{T_n}\right) = \mu \cdot p_z^n \cdot \mathbf{1}$

Computational Bottleneck.

Each of the $|S|$ entries of $\mu \cdot p_z^n$ may be a polynomial of degree up to n

Goal.

Approximate (in **total variation** distance) the distribution of T_n assuming p_z^ℓ, for certain $\ell > 0$, is the largest power of p_z that can be computed explicitly.

[1] Flajolet-Sedgewick'09

[2] Nicodème-Salvy-Flajolet'99, Ll.-Betterton-Knight'08

Approximation based on Doeblin's coefficient.

- **Doeblin's ergodicity coefficient** [1]:

$$\begin{aligned}\alpha(p) &:= \sum_j \min_i p(i,j)\\ &\overset{2}{=} \max\Big\{\alpha \geq 0 \text{ s.t. } (\exists \text{ prob. } e)\,(\forall i,j\in S)\colon\ p(i,j)\geq \alpha\cdot e(j)\Big\}\end{aligned}$$

Remark. Think of e as a stochastic matrix with identical rows

Theorem. [2]

For all $0 \leq \alpha \leq \alpha(p)$ there is a prob. e and stoch. matrix r such that

$$p = \alpha \cdot e + (1-\alpha)\cdot r$$

If $\alpha = \alpha(p) > 0$ then

$$e(j) = \min_i \frac{p(i,j)}{\alpha(p)}$$

[1] Doeblin'37

[2] Chestnut-Ll.'10

Approximation based on Doeblin's coefficient: Toy Example.

$$p \quad := \quad \begin{bmatrix} .3 & .4 & .3 \\ .1 & .2 & .7 \\ .5 & 0 & .5 \end{bmatrix} \qquad \text{i.e.} \qquad \alpha(p) = .1 + 0 + .3 = .4$$

$$= \begin{bmatrix} .1 & 0 & .3 \\ .1 & 0 & .3 \\ .1 & 0 & .3 \end{bmatrix} + \begin{bmatrix} .2 & .4 & 0 \\ 0 & .2 & .4 \\ .4 & 0 & .2 \end{bmatrix}$$

$$= 0.4 \cdot \begin{bmatrix} .25 & 0 & .75 \\ .25 & 0 & .75 \\ .25 & 0 & .75 \end{bmatrix} + 0.6 \cdot \begin{bmatrix} .\bar{3} & .\bar{6} & 0 \\ 0 & .\bar{3} & .\bar{6} \\ .\bar{6} & 0 & .\bar{3} \end{bmatrix}$$

Simulation of the Markov chain.

- 40% of the time jump using e i.e. regardless of the current state (!)
- 60% of the time jump using r

Computational Implication

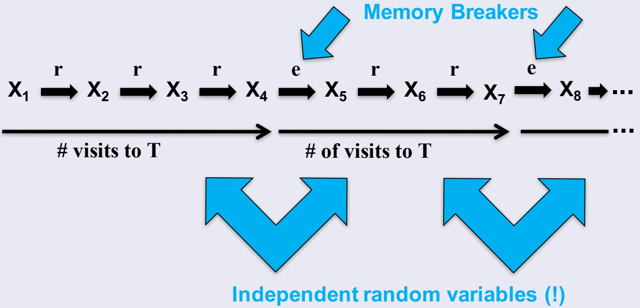

Figure. Conditional on when e and r are used, the sojourn-time of T may be decomposed as a sum of independent random variables i.e.

e acts as a memory breaker for the chain

Longest Head-run in Coin Flips [1]

In n transitions, the longest run of r's concentrates around $\log_{\frac{1}{1-\alpha}}(\alpha\, n)$

[1] Feller'68, Arratia-Goldstein-Gordon'90

Approximation based on Doeblin's coefficient.

Theorem. [1]

Assume $\alpha = \alpha(p) > 0$

If $T_{m,n}$ is the sojourn-time of T in n transitions when we condition the longest run of r to be at most $\ell \sim c \cdot \log_{\frac{1}{1-\alpha}}(n\alpha)$, then

$$\|T_n - T_{\ell,n}\| = O(n^{1-c})$$

Further, the moment generating function of $T_{\ell,n}$ satisfies an explicitly linear recursion of order $(\ell + 1)$

[1] Chestnut-Ll.'10, Ll.-Chestnut'14

WORK IN PROGRESS

(Contact speaker if you would like the complete set of slides!)