Numerical solution of Markov chains and queueing problems

Beatrice Meini

Dipartimento di Matematica, Università di Pisa, Italy

Computational science day, Coimbra, July 23, 2004

Outline

UNIVERSITÀ DI PISA

Outline

1. Introduction to Markov chains
2. Markov chains of M/G/1-type
   - Introduction
   - A power series matrix equation
   - The steady state vector
3. Algorithms for solving the power series matrix equation
   - Functional iterations
   - Cyclic reduction
   - Doubling method
4. Quasi-Birth-Death processes
5. Tree-like stochastic processes
   - Introduction
   - Algorithms

Università di Pisa

Outline

UNIVERSITÀ DI PISA

Outline

UNIVERSITÀ DI PISA

Outline

UNIVERSITÀ DI PISA

Outline

1. Introduction to Markov chains
2. Markov chains of M/G/1-type
   - Introduction
   - A power series matrix equation
   - The steady state vector
3. Algorithms for solving the power series matrix equation
   - Functional iterations
   - Cyclic reduction
   - Doubling method
4. Quasi-Birth-Death processes
5. Tree-like stochastic processes
   - Introduction
   - Algorithms

Università di Pisa

Motivations in Markov chains

- Markov chains: valid tool for modeling problems of the real world (applied probability, queueing models, performance analysis, communication networks, population growth, economic growth, etc.)
- Source of interesting theoretical and computational problems in Numerical Linear Algebra involving either finite or infinite matrices
- Source of very nice structured matrices: almost block Toeplitz, generalized block Hessenberg, multilevel structures.
- People from numerical linear algebra can provide useful tools to the community of applied probabilists and engineers for solving related problems

Bibliography

- M.F. Neuts, *Structured Stochastic Matrices of M/G/1 Type and Their Applications*, Marcel Dekker, New York, 1989.
- G. Latouche, V. Ramaswami, *Introduction to matrix analytic methods in stochastic modeling*, ASA-SIAM Series on Statistics and Applied Probability. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.
- G. W. Stewart, *Introduction to the numerical solution of Markov chains*, Princeton University Press, Princeton, New Jersey, 1994.
- D.A. Bini, G. Latouche, B. Meini, *Numerical Methods for Structured Markov Chains*, Oxford University Press, 2005 (in press)

Announcement

The Fifth International Conference on Matrix Analytic Methods on Stochastic Models (MAM5)
Pisa (Italy), June 21–24, 2005
www.dm.unipi.it/~mam5
Deadline for paper submission: October 2004

Università di Pisa

Introduction to Markov chains

Definition (Stochastic process)

A stochastic process is a family $\{X_t \in E : \quad t \in T\}$ where

- X_t: random variables
- E: state space (denumerable) (e.g. $E = \mathbb{N}$)
- T: time space (denumerable) (e.g. $T = \mathbb{N}$)

Definition (Markov chain)

A Markov chain is a stochastic process $\{X_n\}_{n \in T}$ such that

$$\mathrm{P}[X_{n+1} = i | X_0 = j_0, X_1 = j_1, \dots, X_n = j_n] = \mathrm{P}[X_{n+1} = i | X_n = j_n]$$

Introduction to Markov chains

Definition (Stochastic process)

A stochastic process is a family $\{X_t \in E : \quad t \in T\}$ where

- X_t: random variables
- E: state space (denumerable) (e.g. $E = \mathbb{N}$)
- T: time space (denumerable) (e.g. $T = \mathbb{N}$)

Definition (Markov chain)

A Markov chain is a stochastic process $\{X_n\}_{n \in T}$ such that

$$\mathrm{P}[X_{n+1} = i | X_0 = j_0, X_1 = j_1, \ldots, X_n = j_n] = \\ \mathrm{P}[X_{n+1} = i | X_n = j_n]$$

- The state X_{n+1} of the system at time $n+1$ depends only on the state X_n at time n. It does not depend on the past history of the system
- Homogeneity assumption: $\mathrm{P}[X_{n+1}=i|X_n=j]=\mathrm{P}[X_1=i|X_0=j] \quad \forall \quad n$
- Transition matrix of the Markov chain

$$P=(p_{i,j})_{i,j\in T}, \qquad p_{i,j}=\mathrm{P}[X_1=j|X_0=i].$$

- P is row-stochastic: $p_{i,j}\geq 0$, $\sum_{j\in T} p_{i,j}=1$.

Università di Pisa

- The state X_{n+1} of the system at time $n+1$ depends only on the state X_n at time n. It does not depend on the past history of the system
- Homogeneity assumption:
  $\mathrm{P}[X_{n+1}=i|X_n=j]=\mathrm{P}[X_1=i|X_0=j] \;\; \forall \;\; n$
- Transition matrix of the Markov chain

$$P=(p_{i,j})_{i,j\in T}, \qquad p_{i,j}=\mathrm{P}[X_1=j|X_0=i].$$

- P is row-stochastic: $p_{i,j}\geq 0$, $\sum_{j\in T} p_{i,j}=1$.

Università di Pisa

- The state X_{n+1} of the system at time $n+1$ depends only on the state X_n at time n. It does not depend on the past history of the system
- Homogeneity assumption:
  $\mathrm{P}[X_{n+1} = i | X_n = j] = \mathrm{P}[X_1 = i | X_0 = j] \quad \forall \quad n$
- Transition matrix of the Markov chain

$$P = (p_{i,j})_{i,j \in T}, \qquad p_{i,j} = \mathrm{P}[X_1 = j | X_0 = i].$$

- P is row-stochastic: $p_{i,j} \geq 0$, $\sum_{j \in T} p_{i,j} = 1$.

Università di Pisa

- The state X_{n+1} of the system at time $n+1$ depends only on the state X_n at time n. It does not depend on the past history of the system
- Homogeneity assumption:
  $\mathrm{P}[X_{n+1} = i | X_n = j] = \mathrm{P}[X_1 = i | X_0 = j] \ \ \forall \ \ n$
- Transition matrix of the Markov chain

$$P = (p_{i,j})_{i,j \in T}, \qquad p_{i,j} = \mathrm{P}[X_1 = j | X_0 = i].$$

- P is row-stochastic: $p_{i,j} \geq 0$, $\sum_{j \in T} p_{i,j} = 1$.

UNIVERSITÀ DI PISA

Status of the system

- Let $\mathbf{x}^{(n)} = (x_i^{(n)})$, where

$$x_i^{(n)} = \mathrm{P}[X_n = i], \quad i = 0, 1, 2, \cdots .$$

$\mathbf{x}^{(n)}$ describes the status of the system at time n (say, probability that at time n there are n customers in the queue)

- From the composition lows of probability it follows that

$$x_i^{(n)} \geq 0$$

$$\|\mathbf{x}^{(n)}\|_1 = \sum_{i=0}^{\infty} x_i^{(n)} = 1$$

$$\mathbf{x}^{(n+1)T} = \mathbf{x}^{(n)T} P$$

- Great interest for $\pi = \lim_n \mathbf{x}^{(n)}$ (if it exists): π represents the asymptotic behaviour of the system as the time grows.

Status of the system

- Let $\mathbf{x}^{(n)} = (x_i^{(n)})$, where

$$x_i^{(n)} = \mathrm{P}[X_n = i], \quad i = 0, 1, 2, \cdots .$$

$\mathbf{x}^{(n)}$ describes the status of the system at time n (say, probability that at time n there are n customers in the queue)

- From the composition lows of probability it follows that

$$x_i^{(n)} \geq 0$$

$$\|\mathbf{x}^{(n)}\|_1 = \sum_{i=0}^{\infty} x_i^{(n)} = 1$$

$$\mathbf{x}^{(n+1)T} = \mathbf{x}^{(n)T} P$$

UNIVERSITÀ DI PISA

- Great interest for $\pi = \lim_n \mathbf{x}^{(n)}$ (if it exists): π represents the asymptotic behaviour of the system as the time grows.

Status of the system

- Let $\mathbf{x}^{(n)} = (x_i^{(n)})$, where

$$x_i^{(n)} = \mathrm{P}[X_n = i], \quad i = 0, 1, 2, \cdots .$$

  $\mathbf{x}^{(n)}$ describes the status of the system at time n (say, probability that at time n there are n customers in the queue)
- From the composition lows of probability it follows that

$$x_i^{(n)} \geq 0$$

$$\|\mathbf{x}^{(n)}\|_1 = \sum_{i=0}^{\infty} x_i^{(n)} = 1$$

$$\mathbf{x}^{(n+1)T} = \mathbf{x}^{(n)T} P$$

- Great interest for $\boldsymbol{\pi} = \lim_n \mathbf{x}^{(n)}$ (if it exists): $\boldsymbol{\pi}$ represents the asymptotic behaviour of the system as the time grows.

UNIVERSITÀ DI PISA

Classification of the states

- A state i is called recurrent if, once the Markov chain has visited state i, it will return to it over and over again.
  - A state i is positive recurrent if the expected return time to state i is finite;
  - it is null recurrent if the expected return time is infinite.
- A state i is called transient if it is not recurrent.
- A state i has periodicity $\delta > 1$ if $P[X_n = i | X_0 = i] > 0$ only if $n = 0 \mod \delta$.

If P is irreducible then all the states are transient, or positive recurrent, or null recurrent.

Classification of the states

- A state i is called recurrent if, once the Markov chain has visited state i, it will return to it over and over again.
  - A state i is positive recurrent if the expected return time to state i is finite;
  - it is null recurrent if the expected return time is infinite.
- A state i is called transient if it is not recurrent.
- A state i has periodicity $\delta > 1$ if $\mathrm{P}[X_n = i | X_0 = i] > 0$ only if $n = 0 \mod \delta$.

If P is irreducible then all the states are transient, or positive recurrent, or null recurrent.

UNIVERSITÀ DI PISA

Classification of the states

- A state i is called recurrent if, once the Markov chain has visited state i, it will return to it over and over again.
  - A state i is positive recurrent if the expected return time to state i is finite;
  - it is null recurrent if the expected return time is infinite.
- A state i is called transient if it is not recurrent.
- A state i has periodicity $\delta > 1$ if $\mathrm{P}[X_n = i | X_0 = i] > 0$ only if $n = 0 \mod \delta$.

If P is irreducible then all the states are transient, or positive recurrent, or null recurrent.

UNIVERSITÀ DI PISA

Classification of the states

- A state i is called recurrent if, once the Markov chain has visited state i, it will return to it over and over again.
    - A state i is positive recurrent if the expected return time to state i is finite;
    - it is null recurrent if the expected return time is infinite.
- A state i is called transient if it is not recurrent.
- A state i has periodicity $\delta > 1$ if $\mathrm{P}[X_n = i | X_0 = i] > 0$ only if $n = 0 \mod \delta$.

If P is irreducible then all the states are transient, or positive recurrent, or null recurrent.

UNIVERSITÀ DI PISA

Classification of the states

- A state i is called recurrent if, once the Markov chain has visited state i, it will return to it over and over again.
  - A state i is positive recurrent if the expected return time to state i is finite;
  - it is null recurrent if the expected return time is infinite.
- A state i is called transient if it is not recurrent.
- A state i has periodicity $\delta > 1$ if $\mathrm{P}[X_n = i | X_0 = i] > 0$ only if $n = 0 \mod \delta$.

If P is irreducible then all the states are transient, or positive recurrent, or null recurrent.

UNIVERSITÀ DI PISA

Classification of the states

- A state i is called recurrent if, once the Markov chain has visited state i, it will return to it over and over again.
  - A state i is positive recurrent if the expected return time to state i is finite;
  - it is null recurrent if the expected return time is infinite.
- A state i is called transient if it is not recurrent.
- A state i has periodicity $\delta > 1$ if $\mathrm{P}[X_n = i | X_0 = i] > 0$ only if $n = 0 \mod \delta$.

If P is irreducible then all the states are transient, or positive recurrent, or null recurrent.

Positive recurrence

Theorem

Assume that the Markov chain is irreducible. The states are positive recurrent if and only if there exists a strictly positive invariant probability vector, that is, a vector $\boldsymbol{\pi} = (\pi_i)$ *such that* $\pi_i > 0$ *for all* i*, with*

$$\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T \qquad \text{and} \qquad \sum_i \pi_i = 1.$$

In that case, if the Markov chain is non-periodic, then $\lim_{n\to+\infty} \mathrm{P}[X_n = j | X_0 = i] = \pi_j$ *for all* j*, independently of* i*, and* $\boldsymbol{\pi}$ *is called steady state vector.*

The finite case

For finite matrices the Perron-Frobenius theorem allows to easily give conditions for positive recurrence:

Theorem (Perron-Frobenius)

If $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$, $a_{i,j} \geq 0$ and irreducible then

- $\rho(A) > 0$.
- $\rho(A) > 0$ *is a simple eigenvalue.*
- *If A is non-periodic, then any other eigenvalue λ of A is such that $|\lambda| < \rho(A)$.*
- *There exist unique (up to scaling) positive vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ such that $A\mathbf{x} = \rho(A)\mathbf{x}$, $\mathbf{y}^T A = \rho(A)\mathbf{y}^T$.*

Therefore a finite irreducible Markov chain is positive recurrent

The finite case

For finite matrices the Perron-Frobenius theorem allows to easily give conditions for positive recurrence:

Theorem (Perron-Frobenius)

If $A = (a_{i,j}) \in \mathbb{R}^{n\times n}$, $a_{i,j} \geq 0$ *and irreducible then*

- $\rho(A) > 0$.
- $\rho(A) > 0$ *is a simple eigenvalue.*
- *If* A *is non-periodic, then any other eigenvalue* λ *of* A *is such that* $|\lambda| < \rho(A)$.
- *There exist unique (up to scaling) positive vectors* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *such that* $A\mathbf{x} = \rho(A)\mathbf{x}$, $\mathbf{y}^T A = \rho(A)\mathbf{y}^T$.

Therefore a finite irreducible Markov chain is positive recurrent

The finite case

For finite matrices the Perron-Frobenius theorem allows to easily give conditions for positive recurrence:

Theorem (Perron-Frobenius)

If $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$, $a_{i,j} \geq 0$ *and irreducible then*

- $\rho(A) > 0$.
- $\rho(A) > 0$ *is a simple eigenvalue.*
- *If* A *is non-periodic, then any other eigenvalue* λ *of* A *is such that* $|\lambda| < \rho(A)$.
- *There exist unique (up to scaling) positive vectors* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *such that* $A\mathbf{x} = \rho(A)\mathbf{x}$, $\mathbf{y}^T A = \rho(A)\mathbf{y}^T$.

Therefore a finite irreducible Markov chain is positive recurrent

The infinite case

Let us assume that $P = (p_{i,j})_{i,j\in\mathbb{N}}$ is semi-infinite.
If P is stochastic, the irreducibility of P does not guarantee the existence of a vector $\boldsymbol{\pi} > 0$ such that

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T P, \qquad \|\boldsymbol{\pi}\|_1 = 1.$$

Example

For the stochastic irreducible matrix

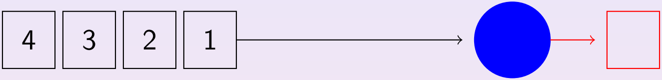
$$P = \begin{bmatrix} 0 & 1 & & & 0 \\ 1/2 & 0 & 1/2 & & \\ & 1/2 & 0 & 1/2 & \\ 0 & & \ddots & \ddots & \ddots \end{bmatrix}$$

one has $\boldsymbol{\pi}^T = \boldsymbol{\pi}^T P$ with $\boldsymbol{\pi}^T = (1/2, 1, 1, 1, \ldots)$

Outline

UNIVERSITÀ DI PISA

A simple queueing problem



- One server which attends to one customer at a time, in order of their arrivals.
- Time is discretized into intervals of fixed length.
- A random number of customers joins the system during each interval.
- Customers are indefinitely patient!

UNIVERSITÀ DI PISA

A simple queueing problem

- Define:
  - α_n: the number of new arrivals in $(n-1, n)$;
  - X_n: the number of customers in the system at time n.
- Then

$$X_{n+1} = \begin{cases} X_n + \alpha_{n+1} - 1 & \text{if } X_n + \alpha_{n+1} \geq 1 \\ 0 & \text{if } X_n + \alpha_{n+1} = 0. \end{cases}$$

- If $\{\alpha_n\}$ are independent random variables, then $\{X_n\}$ is a Markov chain with space state $\mathbb{N}$.
- If in addition the α_n's are identically distributed, then $\{X_n\}$ is homogeneous.

UNIVERSITÀ DI PISA

A simple queueing problem

- Define:
  - α_n: the number of new arrivals in $(n-1, n)$;
  - X_n: the number of customers in the system at time n.
- Then

$$X_{n+1} = \begin{cases} X_n + \alpha_{n+1} - 1 & \text{if } X_n + \alpha_{n+1} \geq 1 \\ 0 & \text{if } X_n + \alpha_{n+1} = 0. \end{cases}$$

- If $\{\alpha_n\}$ are independent random variables, then $\{X_n\}$ is a Markov chain with space state $\mathbb{N}$.
- If in addition the α_n's are identically distributed, then $\{X_n\}$ is homogeneous.

UNIVERSITÀ DI PISA

A simple queueing problem

- Define:
  - α_n: the number of new arrivals in $(n-1, n)$;
  - X_n: the number of customers in the system at time n.
- Then

$$X_{n+1} = \begin{cases} X_n + \alpha_{n+1} - 1 & \text{if } X_n + \alpha_{n+1} \geq 1 \\ 0 & \text{if } X_n + \alpha_{n+1} = 0. \end{cases}$$

- If $\{\alpha_n\}$ are independent random variables, then $\{X_n\}$ is a Markov chain with space state $\mathbb{N}$.
- If in addition the α_n's are identically distributed, then $\{X_n\}$ is homogeneous.

Università di Pisa

A simple queueing problem

The transition matrix $P = (p_{i,j})_{i,j \in \mathbb{N}}$, such that

$$p_{i,j} = \mathrm{P}[X_1 = j | X_0 = i], \qquad \text{for all } i, j \text{ in } \mathbb{N}.$$

is

$$P = \begin{bmatrix} q_0 + q_1 & q_2 & q_3 & \dots \\ q_0 & q_1 & q_2 & \ddots \\ & q_0 & q_1 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix}$$

where q_i is the probability that i new customers join the queue during a unit time interval.

Important families of Markov chains

- M/G/1-type: P is in upper block Hessenberg form, and almost block Toeplitz
- G/M/1-type: P is in lower block Hessenberg form, and almost block Toeplitz
- QBD (Quasi-Birth-Death): P is block tridiagonal, and almost block Toeplitz
- NSF (Non-Skip-Free): P is in generalized block Hessenberg form, and almost block Toeplitz
- Tree-like stohastic process: P has a "recursive structure"

Important families of Markov chains

- M/G/1-type: P is in upper block Hessenberg form, and almost block Toeplitz
- G/M/1-type: P is in lower block Hessenberg form, and almost block Toeplitz
- QBD (Quasi-Birth-Death): P is block tridiagonal, and almost block Toeplitz
- NSF (Non-Skip-Free): P is in generalized block Hessenberg form, and almost block Toeplitz
- Tree-like stohastic process: P has a “recursive structure”

Important families of Markov chains

- M/G/1-type: P is in upper block Hessenberg form, and almost block Toeplitz
- G/M/1-type: P is in lower block Hessenberg form, and almost block Toeplitz
- QBD (Quasi-Birth-Death): P is block tridiagonal, and almost block Toeplitz
- NSF (Non-Skip-Free): P is in generalized block Hessenberg form, and almost block Toeplitz
- Tree-like stohastic process: P has a "recursive structure"

Important families of Markov chains

- M/G/1-type: P is in upper block Hessenberg form, and almost block Toeplitz
- G/M/1-type: P is in lower block Hessenberg form, and almost block Toeplitz
- QBD (Quasi-Birth-Death): P is block tridiagonal, and almost block Toeplitz
- NSF (Non-Skip-Free): P is in generalized block Hessenberg form, and almost block Toeplitz
- Tree-like stohastic process: P has a "recursive structure"

Università di Pisa

Important families of Markov chains

- M/G/1-type: P is in upper block Hessenberg form, and almost block Toeplitz
- G/M/1-type: P is in lower block Hessenberg form, and almost block Toeplitz
- QBD (Quasi-Birth-Death): P is block tridiagonal, and almost block Toeplitz
- NSF (Non-Skip-Free): P is in generalized block Hessenberg form, and almost block Toeplitz
- Tree-like stohastic process: P has a "recursive structure"

M/G/1-type Markov chains

- Introduced by M. F. Neuts in the 80's, they model a large variety of queueing problems.
- The transition matrix is

$$P = \begin{bmatrix} B_0 & B_1 & B_2 & B_3 & \dots \\ A_{-1} & A_0 & A_1 & A_2 & \dots \\ & A_{-1} & A_0 & A_1 & \ddots \\ & & A_{-1} & A_0 & \ddots \\ 0 & & & \ddots & \ddots \end{bmatrix}$$

where $A_{i-1}, B_i \in \mathbb{R}^{m \times m}$, for $i \geq 0$, are nonnegative such that $\sum_{i=-1}^{+\infty} A_i$, $\sum_{i=0}^{+\infty} B_i$, are stochastic.

P is upper block Hessenberg and is block Toeplitz except for its first block row.

Positive recurrence (informal)

Intuitively, positive recurrence means that the global probability that the state changes into a "forward" state is less than the global probability of a change into a "backward" state.
In this way, the probabilities π_i of the stationary probability vector get smaller and smaller as long as i grows.

Example (Positive recurrent Markov chain)

$$P = \begin{bmatrix} 0 & 1 & & & 0 \\ 3/4 & 0 & 1/4 & & \\ & 3/4 & 0 & 1/4 & \\ 0 & & \ddots & \ddots & \ddots \end{bmatrix},$$

$$\boldsymbol{\pi}^T = \left[\frac{1}{2}, \frac{2}{3}, \frac{2}{9}, \frac{2}{27}, \dots\right] \in L^1$$

Transient (informal)

Intuitively, transient means that the global probability that the state changes into a "backward" state is less than the global probability of a change into a "forward" state.

Example (Transient Markov chain)

$$P = \begin{bmatrix} 0 & 1 & & & 0 \\ 1/4 & 0 & 3/4 & & \\ & 1/4 & 0 & 3/4 & \\ 0 & & \ddots & \ddots & \ddots \end{bmatrix},$$

$$\pi^T = [1, 4, 12, 16, \ldots] \notin L^\infty$$

Null recurrence (informal)

Intuitively, null recurrence means that the global probability that the state changes into a "backward" state is equal to the global probability of a change into a "forward" state.

Example (Null recurrence)

$$P = \begin{bmatrix} 0 & 1 & & & 0 \\ 1/2 & 0 & 1/2 & & \\ & 1/2 & 0 & 1/2 & \\ 0 & & \ddots & \ddots & \ddots \end{bmatrix},$$

$$\pi^T = [1/2, 1, 1, \ldots] \notin L^1$$

Positive recurrence

- For M/G/1-type Markov chains positive recurrence is equivalent to

$$\mathbf{b}^T\mathbf{a} < 1,$$

where

$$\mathbf{b}^T = \mathbf{1}^T \sum_{i=1}^{\infty} iA_{i-1}, \qquad \mathbf{1}^T = (1, 1, \ldots, 1),$$

$$\mathbf{a}^T = \mathbf{a}^T \sum_{i=-1}^{\infty} A_i, \quad \mathbf{a}^T\mathbf{1} = 1$$

- Throughout we assume that the Markov chain is irreducible and positive recurrent, therefore there exists the steady state vector $\pi > 0$.

Positive recurrence

- For M/G/1-type Markov chains positive recurrence is equivalent to

$$\mathbf{b}^T\mathbf{a} < 1,$$

where

$$\mathbf{b}^T = \mathbf{1}^T \sum_{i=1}^{\infty} iA_{i-1}, \qquad \mathbf{1}^T = (1, 1, \ldots, 1),$$

$$\mathbf{a}^T = \mathbf{a}^T \sum_{i=-1}^{\infty} A_i, \quad \mathbf{a}^T\mathbf{1} = 1$$

- Throughout we assume that the Markov chain is irreducible, and positive recurrent, therefore there exists the steady state vector $\boldsymbol{\pi} > 0$.

UNIVERSITÀ DI PISA

A power series matrix equation

Theorem (Neuts '98)

The matrix equation

$$X = A_{-1} + A_0 X + A_1 X^2 + A_2 X^3 + \cdots$$

has a minimal component-wise solution G, among the nonnegative solutions.

UNIVERSITÀ DI PISA

Some properties of G

Let $S(z) = zI - \sum_{i=-1}^{+\infty} z^{i+1} A_i$.
If the M/G/1-type Markov chain is positive recurrent, then:

- G is row stochastic.
- $\det S(z)$ has exactly m zeros in the closed unit disk.
- The eigenvalues of G are the zeros of $\det S(z)$ in the closed unit disk.

Therefore G is the minimal solvent (Gohberg, Lancaster, Rodman '82)

UNIVERSITÀ DI PISA

Some properties of G

Let $S(z) = zI - \sum_{i=-1}^{+\infty} z^{i+1} A_i$.
If the M/G/1-type Markov chain is positive recurrent, then:

- G is row stochastic.
- $\det S(z)$ has exactly m zeros in the closed unit disk.
- The eigenvalues of G are the zeros of $\det S(z)$ in the closed unit disk.

Therefore G is the minimal solvent (Gohberg, Lancaster, Rodman '82)

Some properties of $S(z)$

- The power series $S(z) = zI - \sum_{i=-1}^{+\infty} z^{i+1} A_i$ belongs to the Wiener algebra $\mathcal{W}$, therefore it is analytic for $|z| < 1$, continuous for $|z| = 1$.
- Under some mild additional assumptions $S(z)$ is analytic for $|z| < r$, where $r > 1$, and there exists a smallest modulus zero ξ of $\det S(z)$ such that $1 < |\xi| < r$.

UNIVERSITÀ DI PISA

Canonical factorization

Theorem

The function $\phi(z) = I - \sum_{i=-1}^{+\infty} z^i A_i$ *has a weak canonical factorization in* $\mathcal{W}$

$$\phi(z) = \Big(I - \sum_{i=0}^{+\infty} z^i U_i\Big)(I - z^{-1}G), \quad |z| = 1,$$

where:

- $U(z) = I - \sum_{i=0}^{+\infty} z^i U_i$ *is analytic for* $|z| < 1$, $\det U(z) \neq 0$ *for* $|z| \leq 1$;
- $L(z) = I - z^{-1}G$ *is analytic for* $|z| > 1$, $\det L(z) \neq 0$ *for* $|z| > 1$, $\det L(1) = 0$.

Canonical factorization

Theorem

The function $\phi(z) = I - \sum_{i=-1}^{+\infty} z^i A_i$ *has a weak canonical factorization in* $\mathcal{W}$

$$\phi(z) = \Big(I - \sum_{i=0}^{+\infty} z^i U_i\Big)(I - z^{-1}G), \quad |z| = 1,$$

where:

- $U(z) = I - \sum_{i=0}^{+\infty} z^i U_i$ *is analytic for* $|z| < 1$, $\det U(z) \neq 0$ *for* $|z| \leq 1$;
- $L(z) = I - z^{-1}G$ *is analytic for* $|z| > 1$, $\det L(z) \neq 0$ *for* $|z| > 1$, $\det L(1) = 0$.

DI PISA

Canonical factorization

Theorem

The function $\phi(z) = I - \sum_{i=-1}^{+\infty} z^i A_i$ *has a weak canonical factorization in* $\mathcal{W}$

$$\phi(z) = \left(I - \sum_{i=0}^{+\infty} z^i U_i\right)(I - z^{-1}G), \quad |z| = 1,$$

where:

- $U(z) = I - \sum_{i=0}^{+\infty} z^i U_i$ *is analytic for* $|z| < 1$, $\det U(z) \neq 0$ *for* $|z| \leq 1$;
- $L(z) = I - z^{-1}G$ *is analytic for* $|z| > 1$, $\det L(z) \neq 0$ *for* $|z| > 1$, $\det L(1) = 0$.

Matrix interpretation

$$H = \begin{bmatrix} I - A_0 & -A_1 & A_2 & \dots \\ -A_{-1} & I - A_0 & -A_1 & \ddots \\ & -A_{-1} & I - A_0 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} = UL$$

where

$$U = \begin{bmatrix} U_0 & U_1 & U_2 & \dots \\ & U_0 & U_1 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix}, \quad L = \begin{bmatrix} I & & & 0 \\ -G & I & & \\ & -G & I & \\ 0 & & \ddots & \ddots \end{bmatrix}$$

Computing π

Consider the problem of computing $\boldsymbol{\pi} = (\boldsymbol{\pi}_i)_{i\in\mathbb{N}}$, $\boldsymbol{\pi}_i \in \mathbb{R}^m$, such that $\boldsymbol{\pi}^T(I - P) = 0$. i.e.,

$$[\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_1^T, \boldsymbol{\pi}_2^T, \ldots] \left[\begin{array}{c|cccc} I - B_0 & -B_1 & -B_2 & -B_3 & \ldots \\ \hline -A_{-1} & I - A_0 & -A_1 & -A_2 & \ldots \\ & -A_{-1} & I - A_0 & -A_1 & \ddots \\ 0 & & \ddots & \ddots & \ddots \end{array}\right] = 0$$

UNIVERSITÀ DI PISA

Computing π

Then,

$$0 = [\pi_0^T, \pi_1^T, \pi_2^T, \ldots] \left[\begin{array}{c|ccc} I - B_0 & -B_1 & -B_2 & \ldots \\ \hline -A_{-1} & & & \\ 0 & & UL & \\ \vdots & & & \end{array}\right]$$

is equivalent to

$$0 = [\pi_0^T, \pi_1^T, \pi_2^T, \ldots] \left[\begin{array}{c|ccc} I - B_0 & -B_1^* & -B_2^* & \ldots \\ \hline -A_{-1} & & & \\ 0 & & U & \\ \vdots & & & \end{array}\right]$$

where

$$[B_1^*, B_2^*, \ldots] = [B_1, B_1, \ldots]L^{-1}$$

UNIVERSITÀ DI PISA

Computing π

Then,

$$0 = [\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_1^T, \boldsymbol{\pi}_2^T, \ldots] \left[\begin{array}{c|ccc} I - B_0 & -B_1 & -B_2 & \ldots \\ \hline -A_{-1} & & & \\ 0 & & UL & \\ \vdots & & & \end{array}\right]$$

is equivalent to

$$0 = [\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_1^T, \boldsymbol{\pi}_2^T, \ldots] \left[\begin{array}{c|ccc} I - B_0 & -B_1^* & -B_2^* & \ldots \\ \hline -A_{-1} & & & \\ 0 & & U & \\ \vdots & & & \end{array}\right]$$

where

$$[B_1^*, B_2^*, \ldots] = [B_1, B_1, \ldots] L^{-1}$$

Università di Pisa

Computing π

$$0 = [\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_1^T, \boldsymbol{\pi}_2^T, \ldots] \left[\begin{array}{c|cccc} I - B_0 & -B_1^* & -B_2^* & -B_3^* & \ldots \\ \hline -A_{-1} & U_0 & U_1 & U_2 & \ldots \\ & & U_0 & U_1 & \ddots \\ 0 & 0 & & \ddots & \ddots \end{array}\right]$$

The first two equations yield

$$0 = [\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_1^T] \begin{bmatrix} I - B_0 & -B_1^* \\ -A_{-1} & U_0 \end{bmatrix}$$

whence we get

$$\boldsymbol{\pi}_0^T (I - B_0 - B_1^* U_0^{-1} A_{-1}) = 0$$

Computing π

$$0 = [\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_1^T, \boldsymbol{\pi}_2^T, \ldots] \left[\begin{array}{c|cccc} I - B_0 & -B_1^* & -B_2^* & -B_3^* & \ldots \\ \hline -A_{-1} & U_0 & U_1 & U_2 & \ldots \\ & & U_0 & U_1 & \ddots \\ 0 & 0 & & \ddots & \ddots \end{array}\right]$$

From the remaining equations we obtain the block triangular block Toeplitz system

$$[\boldsymbol{\pi}_1^T, \boldsymbol{\pi}_2^T, \ldots] \begin{bmatrix} U_0 & U_1 & U_2 & \ldots \\ & U_0 & U_1 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} = \boldsymbol{\pi}_0^T [B_1^*, B_2^*, \ldots]$$

which can be solved either by means of forward substitution or by FFT-based algorithms.

Ramaswami's formula ('89)

Summing up:

Ramaswami's formula

$$\begin{cases} \pi_0^T(I - B_0 - B_1^* U_0^{-1} A_{-1}) = 0 \\ \pi_1^T = \pi_0^T B_1^* U_0^{-1} \\ \pi_2^T = (\pi_0^T B_2^* - \pi_1^T U_1) U_0^{-1} \\ \pi_i^T = (\pi_0^T B_i^* - \pi_1^T U_{i-1} - \cdots - \pi_{i-1}^T U_1) U_0^{-1} \end{cases}$$

where

$$B_i^* = \sum_{j=i}^{+\infty} B_j G^{j-i}, \quad i = 0, 1, 2, \ldots$$

$$U_0^* = I - \sum_{j=0}^{+\infty} A_j G^j, \quad U_i = \sum_{j=i}^{+\infty} A_j G^{j-i}, \quad i = 1, 2, 3, \ldots.$$

UNIVERSITÀ DI PISA

Computational issues

For the stochasticity of P we have $\lim_i B_i = \lim_i A_i = 0$, so that in floating point computation $B_i \approx 0$ for $i > N$ and the infinite summations turn into finite summations

$$B_i^* = \sum_{j=i}^{\infty} B_j G^{j-i} \approx \sum_{j=i}^{N} B_j G^{j-i}, \quad i = 0, 1, \ldots, N$$

- Compute G.
- Compute B_i^*, U_i, $i = 0, 1, 2, \ldots$ by means of back substitution (Horner's rule) ($O(Nm^3)$ ops)
- Compute the dominant left eigenvector π_0 of an $m \times m$ matrix ($O(m^3)$ ops)
- Computing π_i for $i = 1, 2, \ldots, q$ by solving an $q \times q$ block triangular block Toeplitz system ($O(m^3 q \log q)$ ops)

UNIVERSITÀ DI PISA

Computational issues

For the stochasticity of P we have $\lim_i B_i = \lim_i A_i = 0$, so that in floating point computation $B_i \approx 0$ for $i > N$ and the infinite summations turn into finite summations

$$B_i^* = \sum_{j=i}^{\infty} B_j G^{j-i} \approx \sum_{j=i}^{N} B_j G^{j-i}, \quad i = 0, 1, \dots, N$$

- Compute G.
- Compute B_i^*, U_i, $i = 0, 1, 2, \dots$ by means of back substitution (Horner's rule) ($O(Nm^3)$ ops)
- Compute the dominant left eigenvector π_0 of an $m \times m$ matrix ($O(m^3)$ ops)
- Computing π_i for $i = 1, 2, \dots, q$ by solving an $q \times q$ block triangular block Toeplitz system ($O(m^3 q \log q)$ ops)

UNIVERSITÀ DI PISA

Computational issues

For the stochasticity of P we have $\lim_i B_i = \lim_i A_i = 0$, so that in floating point computation $B_i \approx 0$ for $i > N$ and the infinite summations turn into finite summations

$$B_i^* = \sum_{j=i}^{\infty} B_j G^{j-i} \approx \sum_{j=i}^{N} B_j G^{j-i}, \quad i = 0, 1, \ldots, N$$

- Compute G.
- Compute B_i^*, U_i, $i = 0, 1, 2, \ldots$ by means of back substitution (Horner's rule) ($O(Nm^3)$ ops)
- Compute the dominant left eigenvector π_0 of an $m \times m$ matrix ($O(m^3)$ ops)
- Computing π_i for $i = 1, 2, \ldots, q$ by solving an $q \times q$ block triangular block Toeplitz system ($O(m^3 q \log q)$ ops)

Computational issues

For the stochasticity of P we have $\lim_i B_i = \lim_i A_i = 0$, so that in floating point computation $B_i \approx 0$ for $i > N$ and the infinite summations turn into finite summations

$$B_i^* = \sum_{j=i}^{\infty} B_j G^{j-i} \approx \sum_{j=i}^{N} B_j G^{j-i}, \quad i = 0, 1, \dots, N$$

- Compute G.
- Compute B_i^*, U_i, $i = 0, 1, 2, \dots$ by means of back substitution (Horner's rule) ($O(Nm^3)$ ops)
- Compute the dominant left eigenvector π_0 of an $m \times m$ matrix ($O(m^3)$ ops)
- Computing π_i for $i = 1, 2, \dots, q$ by solving an $q \times q$ block triangular block Toeplitz system ($O(m^3 q \log q)$ ops)

UNIVERSITÀ DI PISA

Computational issues

For the stochasticity of P we have $\lim_i B_i = \lim_i A_i = 0$, so that in floating point computation $B_i \approx 0$ for $i > N$ and the infinite summations turn into finite summations

$$B_i^* = \sum_{j=i}^{\infty} B_j G^{j-i} \approx \sum_{j=i}^{N} B_j G^{j-i}, \quad i = 0, 1, \ldots, N$$

- Compute G.
- Compute B_i^*, U_i, $i = 0, 1, 2, \ldots$ by means of back substitution (Horner's rule) ($O(Nm^3)$ ops)
- Compute the dominant left eigenvector π_0 of an $m \times m$ matrix ($O(m^3)$ ops)
- Computing π_i for $i = 1, 2, \ldots, q$ by solving an $q \times q$ block triangular block Toeplitz system ($O(m^3 q \log q)$ ops)

UNIVERSITÀ DI PISA

Outline

UNIVERSITÀ DI PISA

Functional iterations

Natural iteration

$$\begin{cases} X_{n+1} = \sum_{i=-1}^{+\infty} A_i X_n^{i+1}, \quad n \geq 0 \\ X_0 = 0 \end{cases}$$

History Several variants proposed by Neuts ('81, '89), Ramaswami ('88), Latouche ('93), Bai ('97).

Convergence Convergence analysis performed by Meini ('97), Guo ('99). Convergence is linear, and for some problems it may be extremely slow.

UNIVERSITÀ DI PISA

Some fixed point iterations

Natural iteration

$$X_{n+1} = \sum_{i=-1}^{+\infty} A_i X_n^{i+1}, \quad n \geq 0$$

Traditional iteration

$$X_{n+1} = (I - A_0)^{-1}\left(A_{-1} + \sum_{i=1}^{+\infty} A_i X_n^{i+1}\right), \quad n \geq 0$$

Iteration "based on the matrix U"

$$X_{n+1} = \left(I - \sum_{i=0}^{+\infty} A_i X_n^{i}\right)^{-1} A_{-1}, \quad n \geq 0$$

Some fixed point iterations

Natural iteration

$$X_{n+1} = \sum_{i=-1}^{+\infty} A_i X_n^{i+1}, \quad n \geq 0$$

Traditional iteration

$$X_{n+1} = (I - A_0)^{-1}\left(A_{-1} + \sum_{i=1}^{+\infty} A_i X_n^{i+1}\right), \quad n \geq 0$$

Iteration "based on the matrix U"

$$X_{n+1} = \left(I - \sum_{i=0}^{+\infty} A_i X_n^{i}\right)^{-1} A_{-1}, \quad n \geq 0$$

Some fixed point iterations

Natural iteration

$$X_{n+1} = \sum_{i=-1}^{+\infty} A_i X_n^{i+1}, \quad n \geq 0$$

Traditional iteration

$$X_{n+1} = (I - A_0)^{-1}\left(A_{-1} + \sum_{i=1}^{+\infty} A_i X_n^{i+1}\right), \quad n \geq 0$$

Iteration "based on the matrix U"

$$X_{n+1} = \left(I - \sum_{i=0}^{+\infty} A_i X_n^{i}\right)^{-1} A_{-1}, \quad n \geq 0$$

Convergence analysis: case $X_0 = 0$

Theorem (Latouche '91)

If $X_0 = 0$ then the sequences $\{X_n^{(N)}\}_{n\geq 0}$, $\{X_n^{(T)}\}_{n\geq 0}$, $\{X_n^{(U)}\}_{n\geq 0}$ converge monotonically to the matrix G, that is $X_{n+1} - X_n \geq 0$ for X_n being any of $X_n^{(N)}$, $X_n^{(T)}$, $X_n^{(U)}$. Moreover, for any $n \geq 0$, it holds

$$X_n^{(N)} \leq X_n^{(T)} \leq X_n^{(U)}.$$

Therefore the sequence $\{X_n^{(U)}\}_{n\geq 0}$ provides the best approximation.

Has it the fastest convergence?

Convergence analysis: case $X_0 = 0$

Theorem (Latouche '91)

If $X_0 = 0$ then the sequences $\{X_n^{(N)}\}_{n\geq 0}$, $\{X_n^{(T)}\}_{n\geq 0}$, $\{X_n^{(U)}\}_{n\geq 0}$ converge monotonically to the matrix G, that is $X_{n+1} - X_n \geq 0$ for X_n being any of $X_n^{(N)}$, $X_n^{(T)}$, $X_n^{(U)}$. Moreover, for any $n \geq 0$, it holds

$$X_n^{(N)} \leq X_n^{(T)} \leq X_n^{(U)}.$$

Therefore the sequence $\{X_n^{(U)}\}_{n\geq 0}$ provides the best approximation.

Has it the fastest convergence?

Convergence analysis: case $X_0 = 0$

Consider for simplicity the natural iteration.
Define $E_n = G - X_n$ the error at step n.

Theorem

1. $0 \le E_{n+1} \le E_n$ *for any* $n \ge 0$.
2. $E_{n+1}\mathbf{1} = R_n E_n \mathbf{1}$ *where* $R_n = \sum_{i=0}^{+\infty} \sum_{j=i}^{+\infty} A_j X_n^{j-i}$.
3. $\|E_n\|_\infty = \left\| \prod_{i=0}^{n-1} R_i \right\|_\infty$.

Denoting $r = \lim_n \sqrt[n]{\|E_n\|}$, *one has* $r = \rho(R)$, *where*

$$R = \lim_n R_n = \sum_{i=0}^{+\infty} \sum_{j=i}^{+\infty} A_j G^{j-i}.$$

4.

Convergence analysis: case $X_0 = 0$

Consider for simplicity the natural iteration.
Define $E_n = G - X_n$ the error at step n.

Theorem

1. $0 \le E_{n+1} \le E_n$ *for any* $n \ge 0$.
2. $E_{n+1}\mathbf{1} = R_n E_n \mathbf{1}$ *where* $R_n = \sum_{i=0}^{+\infty} \sum_{j=i}^{+\infty} A_j X_n^{j-i}$.
3. $\|E_n\|_\infty = \left\| \prod_{i=0}^{n-1} R_i \right\|_\infty$.
4.

Denoting $r = \lim_n \sqrt[n]{\|E_n\|}$, *one has* $r = \rho(R)$, *where*

$$R = \lim_n R_n = \sum_{i=0}^{+\infty} \sum_{j=i}^{+\infty} A_j G^{j-i}.$$

Convergence analysis: case $X_0 = 0$

Consider for simplicity the natural iteration.
Define $E_n = G - X_n$ the error at step n.

Theorem

1. $0 \le E_{n+1} \le E_n$ *for any* $n \ge 0$.
2. $E_{n+1}\mathbf{1} = R_n E_n \mathbf{1}$ *where* $R_n = \sum_{i=0}^{+\infty} \sum_{j=i}^{+\infty} A_j X_n^{j-i}$.
3. $\|E_n\|_\infty = \left\| \prod_{i=0}^{n-1} R_i \right\|_\infty$.

Denoting $r = \lim_n \sqrt[n]{\|E_n\|}$, *one has* $r = \rho(R)$, *where*

$$R = \lim_n R_n = \sum_{i=0}^{+\infty} \sum_{j=i}^{+\infty} A_j G^{j-i}.$$

4.

Convergence analysis: case $X_0 = 0$

Consider for simplicity the natural iteration.
Define $E_n = G - X_n$ the error at step n.

Theorem

1. $0 \le E_{n+1} \le E_n$ *for any* $n \ge 0$.
2. $E_{n+1}\mathbf{1} = R_n E_n \mathbf{1}$ *where* $R_n = \sum_{i=0}^{+\infty}\sum_{j=i}^{+\infty} A_j X_n^{j-i}$.
3. $\|E_n\|_\infty = \left\| \prod_{i=0}^{n-1} R_i \right\|_\infty$.
4. *Denoting* $r = \lim_n \sqrt[n]{\|E_n\|}$, *one has* $r = \rho(R)$, *where*

$$R = \lim_n R_n = \sum_{i=0}^{+\infty}\sum_{j=i}^{+\infty} A_j G^{j-i}.$$

Convergence analysis: case $X_0 = 0$

Consider for simplicity the natural iteration.

Define $E_n = G - X_n$ the error at step n.

Theorem

1. $0 \le E_{n+1} \le E_n$ *for any* $n \ge 0$.
2. $E_{n+1}\mathbf{1} = R_n E_n \mathbf{1}$ *where* $R_n = \sum_{i=0}^{+\infty} \sum_{j=i}^{+\infty} A_j X_n^{j-i}$.
3. $\|E_n\|_\infty = \left\| \prod_{i=0}^{n-1} R_i \right\|_\infty$.

Denoting $r = \lim_n \sqrt[n]{\|E_n\|}$, *one has* $r = \rho(R)$, *where*

$$R = \lim_n R_n = \sum_{i=0}^{+\infty} \sum_{j=i}^{+\infty} A_j G^{j-i}.$$

Comparison among the 3 iterations

Theorem

One has $r_N = \rho(R^{(N)})$, $r_T = \rho(R^{(T)})$, $r_U = \rho(R^{(U)})$, *where*

$$R^{(N)} = \sum_{i=0}^{+\infty} A_i^*,$$

$$R^{(T)} = (I - A_0)^{-1}\left(\sum_{i=0}^{+\infty} A_i^* - A_0\right),$$

$$R^{(U)} = (I - A_0^*)^{-1}\sum_{i=1}^{+\infty} A_i^*.$$

and

$$0 \le R^{(U)} \le R^{(T)} \le R^{(N)}$$

Convergence analysis: case $X_0 = I$

Consider for simplicity the natural iteration.

Theorem

Under mild irreducibility assumptions, for the convergence rate

$$r = \lim_n \sqrt[n]{\|E_n\|}$$

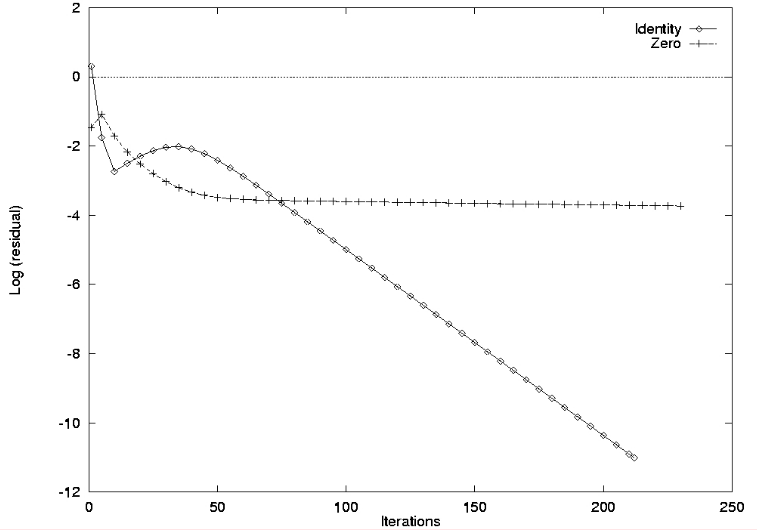
of the sequences obtained with $X_0 = I$, we have

$$r_N = \rho_2(R^{(N)}), \quad r_T = \rho_2(R^{(T)}), \quad r_U = \rho_2(R^{(U)}),$$

where ρ_2 denotes the second largest modulus eigenvalue.

Starting with $X_0 = I$ the convergence is faster

Convergence for $X_0 = 0$ and $X_0 = I$



Università di Pisa

Linearization of the matrix equation

$$\begin{bmatrix} I - A_0 & -A_1 & -A_2 & \cdots \\ -A_{-1} & I - A_0 & -A_1 & \ddots \\ & -A_{-1} & I - A_0 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} G \\ G^2 \\ G^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ 0 \\ \vdots \end{bmatrix}.$$

G can be interpreted by means of the solution of an infinite block Hessenberg, block Toeplitz system

Cyclic reduction: history

- Introduced in the late '60s by Buzbee, Golub and Nielson for solving block tridiagonal systems in the context of elliptic equations.
- Stability and convergence properties: Amodio and Mazzia ('94), Yalamov ('95), Yalamov and Pavlov ('96), etc.
- Rediscovered by Latouche and Ramaswami (Logarithmic reduction) in the context of Markov chains ('93);
- Extended to infinite block Hessenberg, block Toeplitz systems by Bini and Meini (starting from '96).

The cyclic reduction algorithm

Original system:

$$\begin{bmatrix} I-A_0 & -A_1 & -A_2 & \dots \\ -A_{-1} & I-A_0 & -A_1 & \ddots \\ & -A_{-1} & I-A_0 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} G \\ G^2 \\ G^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

The cyclic reduction algorithm

Block even-odd permutation:

$$\left[\begin{array}{ccc|ccc} I-A_0 & -A_2 & \dots & -A_{-1} & -A_1 & \dots \\ & I-A_0 & \ddots & & -A_{-1} & \ddots \\ 0 & & \ddots & 0 & & \ddots \\ \hline -A_1 & -A_3 & \dots & I-A_0 & -A_2 & \dots \\ -A_{-1} & -A_1 & \ddots & & I-A_0 & \ddots \\ 0 & \ddots & \ddots & 0 & & \ddots \end{array}\right] \left[\begin{array}{c} G^2 \\ G^4 \\ \vdots \\ \hline G \\ G^3 \\ \vdots \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ \hline A_{-1} \\ 0 \\ \vdots \end{array}\right]$$

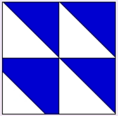
In compact form:

$$\begin{bmatrix} I-H_1 & -H_2 \\ -H_3 & I-H_4 \end{bmatrix} \begin{bmatrix} \mathbf{g}_- \\ \mathbf{g}_+ \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{b} \end{bmatrix}$$

The cyclic reduction algorithm

Structure of the matrix:

$$\begin{bmatrix} I - H_1 & -H_2 \\ -H_3 & I - H_4 \end{bmatrix} =$$

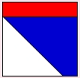


Schur complementation:

$$I - H_4 - H_3(I - H_1)^{-1}H_2 =$$

$+$

$=$



UNIVERSITÀ DI PISA

Upper block Hessenberg matrix, block Toeplitz except for its first block row

The cyclic reduction algorithm

Resulting system:

$$\begin{bmatrix} I-\widehat{A}_0^{(1)} & -\widehat{A}_1^{(1)} & -\widehat{A}_2^{(1)} & \dots \\ -A_{-1}^{(1)} & I-A_0^{(1)} & -A_1^{(1)} & \dots \\ & -A_{-1}^{(1)} & I-A_0^{(1)} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} G \\ G^3 \\ G^5 \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

The cyclic reduction algorithm

One more step of the same procedure:

$$\begin{bmatrix} I-\widehat{A}_0^{(2)} & -\widehat{A}_1^{(2)} & -\widehat{A}_2^{(2)} & \dots \\ -A_{-1}^{(2)} & I-A_0^{(2)} & -A_1^{(2)} & \dots \\ & -A_{-1}^{(2)} & I-A_0^{(2)} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} G \\ G^5 \\ G^9 \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

Università di Pisa

The cyclic reduction algorithm

At the n-th step:

$$\begin{bmatrix} I-\widehat{A}_0^{(n)} & -\widehat{A}_1^{(n)} & -\widehat{A}_2^{(n)} & \dots \\ -A_{-1}^{(n)} & I-A_0^{(n)} & -A_1^{(n)} & \dots \\ & -A_{-1}^{(n)} & I-A_0^{(n)} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} G \\ G^{2^n+1} \\ G^{2\cdot 2^n+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

UNIVERSITÀ DI PISA

The cyclic reduction algorithm

At the limit as $n \to \infty$:

$$\begin{bmatrix} I - \widehat{A}_0^{(\infty)} & & & 0 \\ -A_{-1}^{(\infty)} & I - A_0^{(\infty)} & & \\ & -A_{-1}^{(\infty)} & I - A_0^{(\infty)} & \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} G \\ G^* \\ G^* \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

where $G^* = \lim_n G^n$.
Therefore $G = (I - \widehat{A}_0^{(\infty)})^{-1} A_{-1}$

The cyclic reduction algorithm

Functional interpretation

$$A^{(n+1)}(z) = zA^{(n)}_{\text{odd}}(z) + A^{(n)}_{\text{even}}(z)(I - A^{(n)}_{\text{odd}}(z))^{-1}A^{(n)}_{\text{even}}(z)$$

$$\widehat{A}^{(n+1)}(z) = \widehat{A}^{(n)}_{\text{even}}(z) + \widehat{A}^{(n)}_{\text{odd}}(z)(I - A^{(n)}_{\text{odd}}(z))^{-1}A^{(n)}_{\text{even}}(z)$$

where

$$\widehat{A}^{(n)}(z) = \sum_{i=0}^{+\infty} z^i \widehat{A}^{(n)}_i, \quad A^{(n)}(z) = \sum_{i=-1}^{+\infty} z^{i+1} A^{(n)}_i$$

Applicability of CR: the role of Wiener algebra

Theorem

For any $n \geq 0$ one has:

1. $A^{(n)}(z)$ *and* $\widehat{A}^{(n)}(z)$ *belong to* $\mathcal{W}_+$.
2. $I - A^{(n)}_{\text{odd}}(z)$ *is invertible for* $|z| \leq 1$ *and its inverse belongs to* $\mathcal{W}_+$.
3. $\phi^{(n)}(z) = I - z^{-1}A^{(n)}(z)$ *has a* weak *canonical factorization*

$$\phi^{(n)}(z) = \Big(I - \sum_{i=0}^{+\infty} z^i U_i^{(n)}\Big)(I - z^{-1}G^{2^n}), \quad |z| = 1.$$

Convergence of CR

Theorem

Let ξ *be the zero of smallest modulus of* $\det S(z)$ *such that* $|\xi| > 1$. *Then:*

1. $\{A^{(n)}(z)\}_n \longrightarrow A_{-1}^{(\infty)} + zA_0^{(\infty)}$ *uniformly over any compact subset of* $\{z \in \mathbb{C} : \quad |z| < \xi\}$.
2. $\|A_i^{(n)}\| \le \gamma|\xi|^{-i \cdot 2^n}$ *and* $\|\widehat{A}_i^{(n)}\| \le \gamma|\xi|^{-i \cdot 2^n}$, *for any* $i \ge 1$, $n \ge 0$.
3. $\|\widehat{A}_0^{(n)} - \widehat{A}_0^{(\infty)}\| \le \gamma|\xi|^{-2^n}$ *for any* $n \ge 0$.
4. $\rho(\widehat{A}_0^{(\infty)}) \le \rho(A_0^{(\infty)}) < 1$.
5. $\|G - G^{(n)}\| \le \gamma|\xi|^{-2^n}$, *where* $G^{(n)} = (I - \widehat{A}_0^{(n)})^{-1}A_{-1}$.

Computational issues

The matrix power series $A^{(n)}(z)$, $\widehat{A}^{(n)}(z)$ are approximated by matrix polynomials of degree at most d_n.

The computation of such matrix polynomials by means of evaluation/iterpolation at the roots of unity can be performed in

$$O(m^3 d_n + m^2 d_n \log d_n)$$

arithmetic operations

Doubling method

History Introduced by W.J. Stewart ('95) to solve general block Hessenberg systems, applied by Latouche and Stewart ('95) for computing G, improved by Bini and Meini ('98) by exploiting the Toeplitz structure of the block Hessenberg matrices.

Idea Successively solve finite block Hessenberg systems of block size which doubles at each iterative step.

UNIVERSITÀ DI PISA

Doubling method

Truncation at block size n of the infinite system:

$$\begin{bmatrix} I-A_0 & -A_1 & -A_2 & \dots & -A_{n-1} \\ -A_{-1} & I-A_0 & -A_1 & \ddots & \vdots \\ & -A_{-1} & I-A_0 & \ddots & -A_2 \\ & & \ddots & \ddots & -A_1 \\ 0 & & & -A_{-1} & I-A_0 \end{bmatrix} \begin{bmatrix} X_1^{(n)} \\ X_2^{(n)} \\ \vdots \\ X_n^{(n)} \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Doubling method: convergence

Theorem

For any $n \geq 1$ one has:

- $0 \leq X_1^{(n)} \leq X_1^{(n+1)} \leq G$.
- $X_i^{(n)} \leq G^i$ *for* $i = 1, \ldots, n$.
- *For any $\epsilon > 0$ there exist positive constants γ and σ such that*

$$\|G - X_1^{(n)}\|_\infty \leq \gamma(|\xi| - \epsilon)^{-n},$$

where ξ is the zero of smallest modulus of $\det S(z)$ such that $|\xi| > 1$.

Doubling method: algorithm

The algorithm consists in successively solving systems of block size 2, 4, 8, 16, ...

- Size doubling at each step $\Longrightarrow$ Quadratic convergence
- Use of FFT and Toeplitz structure $\Longrightarrow$ The $2^n \times 2^n$ block system can be solved in $O(m^3 2^n + m^2 n 2^n)$ arithmetic operations.

Università di Pisa

Outline

UNIVERSITÀ DI PISA

Quasi-Birth-Death processes

If $A_i = 0$ for $i > 1$ the M/G/1-type Markov chain is called a Quasi-Birth-Death process (QBD).

Problem

Computation of the minimal component-wise solution G, among the nonnegative solutions, of

$$X = A_{-1} + A_0 X + A_1 X^2$$

Linearization of the matrix equation

$$\begin{bmatrix} I-A_0 & -A_1 & & 0 \\ -A_{-1} & I-A_0 & -A_1 & \\ & -A_{-1} & I-A_0 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \boxed{G} \\ G^2 \\ G^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ 0 \\ \vdots \end{bmatrix}.$$

G can be interpreted by means of the solution of an infinite block triangular, block Toeplitz system

UNIVERSITÀ DI PISA

Cyclic reduction for QBD's

At the n-th step

$$\begin{bmatrix} I-\widehat{A}_0^{(n)} & -A_1^{(n)} & & 0 \\ -A_{-1}^{(n)} & I-A_0^{(n)} & -A_1^{(n)} & \\ & -A_{-1}^{(n)} & I-A_0^{(n)} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \boxed{G} \\ G^{2^n+1} \\ G^{2\cdot 2^n+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ 0 \\ \vdots \end{bmatrix}.$$

UNIVERSITÀ DI PISA

Cyclic reduction for QBD's

At the limit as $n \to \infty$:

$$\begin{bmatrix} I - \widehat{A}_0^{(\infty)} & & & 0 \\ -A_{-1}^{(\infty)} & I - A_0^{(\infty)} & & \\ & -A_{-1}^{(\infty)} & I - A_0^{(\infty)} & \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} G \\ G^* \\ G^* \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

where $G^* = \lim_n G^n$.
Therefore $G = (I - \widehat{A}_0^{(\infty)})^{-1} A_{-1}$

UNIVERSITÀ DI PISA

Cyclic reduction

Recursive (algebraic) relations

$$
\begin{aligned}
A_{-1}^{(n+1)} &= A_{-1}^{(n)} K^{(n)} A_{-1}^{(n)}, \\
A_{0}^{(n+1)} &= A_{0}^{(n)} + A_{-1}^{(n)} K^{(n)} A_{1}^{(n)} + A_{1}^{(n)} K^{(n)} A_{-1}^{(n)}, \\
A_{1}^{(n+1)} &= A_{1}^{(n)} K^{(n)} A_{1}^{(n)}, \\
\widehat{A}_{0}^{(n+1)} &= \widehat{A}_{0}^{(n)} + A_{1}^{(n)} K^{(n)} A_{-1}^{(n)}, \qquad n \geq 0
\end{aligned}
$$

where $K^{(n)} = (I - A_0^{(n)})^{-1}$.

Università di Pisa

Outline

UNIVERSITÀ DI PISA

Tree-like stochastic processes

Motivation Tree-Like processes are used to model certain queueing problems: single server queues with LIFO service discipline, medium access control protocol with an underlying stack structure, etc. (Latouche, Ramaswami '99)

Assumptions B, A_i and D_i, $i = 1, \ldots, d$, nonnegative $m \times m$ matrices, such that B is sub-stochastic and $B + D_i + A_1 + \cdots + A_d$, $i = 1, \ldots, d$, are stochastic. We set $C = I - B$.

Tree-like stochastic processes

The generator matrix has the form

$$Q = \begin{bmatrix} C_0 & \Lambda_1 & \Lambda_2 & \dots & \Lambda_d \\ V_1 & W & 0 & \dots & 0 \\ V_2 & 0 & W & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ V_d & 0 & \dots & 0 & W \end{bmatrix},$$

where C_0 is an $m \times m$ matrix,

$$\Lambda_i = \begin{bmatrix} A_i & 0 & 0 & \dots \end{bmatrix}, \quad V_i = \begin{bmatrix} D_i \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \quad \text{for } 1 \le i \le d$$

Tree-like processes

The infinite matrix W is recursively defined by

$$W = \begin{bmatrix} C & \Lambda_1 & \Lambda_2 & \dots & \Lambda_d \\ V_1 & W & 0 & \dots & 0 \\ V_2 & 0 & W & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ V_d & 0 & \dots & 0 & W \end{bmatrix}.$$

Università di Pisa

Tree-like processes

Theorem

The matrix W can be factorized as $W = UL$, where

$$U = \begin{bmatrix} S & \Lambda_1 & \Lambda_2 & \dots & \Lambda_d \\ 0 & U & 0 & \dots & 0 \\ 0 & 0 & U & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & U \end{bmatrix}, \quad L = \begin{bmatrix} I & 0 & 0 & \dots & 0 \\ Y_1 & L & 0 & \dots & 0 \\ Y_2 & 0 & L & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ Y_d & 0 & \dots & 0 & L \end{bmatrix}$$

and S is the minimal solution of $X + \sum_{i=1}^{d} A_i X^{-1} D_i = C$.

Consequence: Once the matrix S is known, the stationary probability vector can be computed by using the UL factorization of W.

Tree-like processes

Theorem

The matrix W can be factorized as $W = UL$, where

$$U = \begin{bmatrix} S & \Lambda_1 & \Lambda_2 & \dots & \Lambda_d \\ 0 & U & 0 & \dots & 0 \\ 0 & 0 & U & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & U \end{bmatrix}, \quad L = \begin{bmatrix} I & 0 & 0 & \dots & 0 \\ Y_1 & L & 0 & \dots & 0 \\ Y_2 & 0 & L & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ Y_d & 0 & \dots & 0 & L \end{bmatrix}$$

and S is the minimal solution of $X + \sum_{i=1}^{d} A_i X^{-1} D_i = C$.

Consequence: Once the matrix S is known, the stationary probability vector can be computed by using the UL factorization of W.

Natural fixed point iteration

The sequences

$$\begin{cases} S_n = C + \sum_{1\le i\le d} A_i G_{i,n}, \\ G_{i,n+1} = (-S_n)^{-1} D_i, \qquad \text{for } 1 \le i \le d,\ n \ge 0, \end{cases}$$

with $G_{1,0} = \ldots = G_{d,0} = 0$, monotonically converge to S and $G_i = (-S)^{-1} D_i$, $i = 1, \ldots, d$, respectively (Latouche and Ramaswami '99)

UNIVERSITÀ DI PISA

Cyclic reduction + fixed point iteration

- Multiply

$$S + \sum_{j=1}^{d} A_j S^{-1} D_j = C$$

by $S^{-1} D_i$, for $i = 1, \ldots, d$.

- Observe that $G_i = (-S)^{-1} D_i$, $i = 1, \ldots, d$, is a solution

$$D_i + (C + \sum_{\substack{1 \le j \le d \\ j \ne i}} A_j G_j) X + A_i X^2 = 0.$$

- We may prove that G_i is the minimal solvent.

UNIVERSITÀ DI PISA

Cyclic reduction + fixed point iteration

- Multiply

$$S + \sum_{j=1}^{d} A_j S^{-1} D_j = C$$

by $S^{-1} D_i$, for $i = 1, \ldots, d$.

- Observe that $G_i = (-S)^{-1} D_i$, $i = 1, \ldots, d$, is a solution

$$D_i + (C + \sum_{\substack{1 \le j \le d \\ j \ne i}} A_j G_j) X + A_i X^2 = 0.$$

- We may prove that G_i is the minimal solvent.

Università di Pisa

Cyclic reduction + fixed point iteration

- Multiply

$$S + \sum_{j=1}^{d} A_j S^{-1} D_j = C$$

by $S^{-1}D_i$, for $i = 1, \ldots, d$.

- Observe that $G_i = (-S)^{-1}D_i$, $i = 1, \ldots, d$, is a solution

$$D_i + (C + \sum_{\substack{1 \le j \le d \\ j \ne i}} A_j G_j)X + A_i X^2 = 0.$$

- We may prove that G_i is the minimal solvent.

UNIVERSITÀ DI PISA

Cyclic reduction + fixed point iteration

- Set $G_{1,0} = G_{2,0} = \cdots = G_{d,0} = 0$
- For $n = 0, 1, 2, \ldots$
  - For $i = 1, \ldots, d$:
    1. define
    $$F_{i,n} = C + \sum_{1 \le j \le i-1} A_j G_{j,n} + \sum_{i+1 \le j \le d} A_j G_{j,n-1}$$
    2. compute, by means of cyclic reduction, the minimal solvent $G_{i,n}$ of
    $$D_i + F_{i,n} X + A_i X^2 = 0.$$

UNIVERSITÀ DI PISA

The sequences $\{G_{i,n} : n \ge 0\}$ monotonically converge to G_i, for $1 \le i \le d$

Cyclic reduction + fixed point iteration

- Set $G_{1,0} = G_{2,0} = \cdots = G_{d,0} = 0$
- For $n = 0, 1, 2, \ldots$
  - For $i = 1, \ldots, d$:
    1. define

$$F_{i,n} = C + \sum_{1 \le j \le i-1} A_j G_{j,n} + \sum_{i+1 \le j \le d} A_j G_{j,n-1}.$$

    2. compute, by means of cyclic reduction, the minimal solvent $G_{i,n}$ of

$$D_i + F_{i,n} X + A_i X^2 = 0,$$

Università di Pisa

The sequences $\{G_{i,n} : n \ge 0\}$ monotonically converge to G_i, for $1 \le i \le d$

Cyclic reduction + fixed point iteration

- Set $G_{1,0} = G_{2,0} = \cdots = G_{d,0} = 0$
- For $n = 0, 1, 2, \ldots$
  - For $i = 1, \ldots, d$:
    1. define
       $$F_{i,n} = C + \sum_{1 \le j \le i-1} A_j G_{j,n} + \sum_{i+1 \le j \le d} A_j G_{j,n-1}.$$
    2. compute, by means of cyclic reduction, the minimal solvent $G_{i,n}$ of
       $$D_i + F_{i,n} X + A_i X^2 = 0,$$

UNIVERSITÀ DI PISA

The sequences $\{G_{i,n} : n \ge 0\}$ monotonically converge to G_i, for $1 \le i \le d$

Newton's iteration

- Set $S_0 = C$
- For $n = 0, 1, 2, \ldots$
  1. Compute $L_n = S_n - C + \sum_{i=1}^{d} A_i S_n^{-1} D_i$.
  2. Compute the solution Y_n of

$$X - \sum_{i=1}^{d} A_i S_n^{-1} X S_n^{-1} D_i = L_n \qquad (1)$$

  3. Set $S_{n+1} = S_n - Y_n$

The sequence $\{S_n\}_n$ converges quadratically to S.

Open issues: efficient computation of the solution of (1).

Università di Pisa

Newton's iteration

- Set $S_0 = C$
- For $n = 0, 1, 2, \ldots$
  1. Compute $L_n = S_n - C + \sum_{i=1}^{d} A_i S_n^{-1} D_i$.
  2. Compute the solution Y_n of

$$X - \sum_{i=1}^{d} A_i S_n^{-1} X S_n^{-1} D_i = L_n \qquad (1)$$

  3. Set $S_{n+1} = S_n - Y_n$

The sequence $\{S_n\}_n$ converges quadratically to S.

Open issues: efficient computation of the solution of (1).

UNIVERSITÀ DI PISA

Newton's iteration

- Set $S_0 = C$
- For $n = 0, 1, 2, \ldots$
  1. Compute $L_n = S_n - C + \sum_{i=1}^{d} A_i S_n^{-1} D_i$.
  2. Compute the solution Y_n of

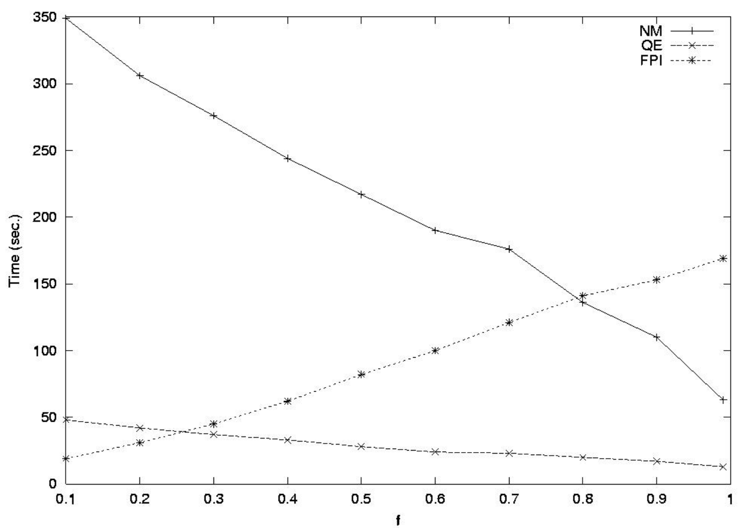
$$X - \sum_{i=1}^{d} A_i S_n^{-1} X S_n^{-1} D_i = L_n \qquad (1)$$

  3. Set $S_{n+1} = S_n - Y_n$

The sequence $\{S_n\}_n$ converges quadratically to S.

Open issues: efficient computation of the solution of (1).

Cpu time



UNIVERSITÀ DI PISA

Wiener algebra

Definition

The Wiener algebra $\mathcal{W}$ is the set of complex $m \times m$ matrix valued functions $A(z) = \sum_{i=-\infty}^{+\infty} z^i A_i$ such that $\sum_{i=-\infty}^{+\infty} |A_i|$ is finite.

Definition

The set $\mathcal{W}_+$ is the subalgebra of $\mathcal{W}$ made up by power series of the kind $\sum_{i=0}^{+\infty} z^i A_i$.

M/G/1 Markov chain

$$P = \begin{bmatrix} B_0 & B_1 & B_2 & B_3 & \dots \\ A_{-1} & A_0 & A_1 & A_2 & \dots \\ & A_{-1} & A_0 & A_1 & \ddots \\ O & & \ddots & \ddots & \ddots \end{bmatrix}$$

$$A_i, B_{i+1} \in \mathbb{R}^{m \times m}, \qquad i = -1, 0, 1, \dots$$

G/M/1 Markov chain

$$P = \begin{bmatrix} B_0 & A_1 & & 0 & \\ B_{-1} & A_0 & A_1 & & \\ B_{-2} & A_{-1} & A_0 & A_1 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}$$

UNIVERSITÀ DI PISA

QBD Stochastic processes

$$P = \begin{bmatrix} B_0 & B_1 & & & & 0 \\ B_{-1} & A_0 & A_1 & & & \\ & A_{-1} & A_0 & A_1 & & \\ & & A_{-1} & A_0 & A_1 & \\ 0 & & & \ddots & \ddots & \ddots \end{bmatrix}$$

Università di Pisa

Non-Skip-Free Stochastic processes

$$P = \begin{bmatrix} B_{0,1} & B_{0,1} & B_{0,2} & B_{0,3} & \dots & \dots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ B_{k-1,0} & B_{k-1,1} & B_{k-1,2} & B_{k-1,3} & \dots & \dots & \\ A_{-k} & A_{-k+1} & A_{-k+2} & A_{-k+3} & \dots & \dots & \\ & A_{-k} & A_{-k+1} & A_{-k+2} & A_{-k+3} & \ddots & \\ & & A_{-k} & A_{-k+1} & A_{-k+2} & A_{-k+3} & \ddots \\ O & & & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Università di Pisa