# Decay properties of certain matrix power series encountered in stochastic processes

Beatrice Meini
joint work with D. Bini and V. Ramaswami

Structured Linear Algebra Problems: Analysis, Algorithms, and Applications
Cortona, September 15–19, 2008

Università di Pisa

# Outline



Università di Pisa

# The problem

- Let $A_n$, for $n \geq -1$, be $k \times k$ nonnegative matrices such that $\sum_{n=-1}^{\infty} A_n$ is stochastic;

# The problem

- Let $A_n$, for $n \geq -1$, be $k \times k$ nonnegative matrices such that $\sum_{n=-1}^{\infty} A_n$ is stochastic;
- For $0 \leq w \leq 1$, let $G(w)$ be the minimal nonnegative solution to the matrix equation

$$X = w \sum_{n=-1}^{\infty} A_n X^{n+1}$$

i.e., for any solution $X(w) \geq 0$, one has $G(w) \leq X(w)$.

Università di Pisa

# The problem

- Let $A_n$, for $n \geq -1$, be $k \times k$ nonnegative matrices such that $\sum_{n=-1}^{\infty} A_n$ is stochastic;
- For $0 \leq w \leq 1$, let $G(w)$ be the minimal nonnegative solution to the matrix equation

$$X = w \sum_{n=-1}^{\infty} A_n X^{n+1}$$

i.e., for any solution $X(w) \geq 0$, one has $G(w) \leq X(w)$.

Our goal: To analyze "decay properties" of $G(w)$ (more details will follow in the next slides)

Università di Pisa

# Motivation

Consider a Markov Renewal Process (MRP) of M/G/1-type with levels $\ell_0, \ell_1, \ell_2, \ldots$, defined by the kernel

$$K(x) = \begin{bmatrix} \widetilde{B}_0(x) & \widetilde{B}_1(x) & \widetilde{B}_2(x) & \widetilde{B}_3(x) & \ldots \\ \widetilde{A}_{-1}(x) & \widetilde{A}_0(x) & \widetilde{A}_1(x) & \widetilde{A}_2(x) & \ldots \\ & \widetilde{A}_{-1}(x) & \widetilde{A}_0(x) & \widetilde{A}_1(x) & \ddots \\ & & \widetilde{A}_{-1}(x) & \widetilde{A}_0(x) & \ddots \\ 0 & & & \ddots & \ddots \end{bmatrix}, \quad x \geq 0,$$

where

$$\widetilde{A}_k(x) = P\{X_{n+1} \in \ell_{i+k},\ \tau_{n+1} - \tau_n \leq x | \quad X_0, \ldots, X_n, \tau_0, \ldots, \tau_n, X_n \in \ell_i\},$$

$$\widetilde{B}_k(x) = P\{X_{n+1} \in \ell_k,\ \tau_{n+1} - \tau_n \leq x | \quad X_0, \ldots, X_n, \tau_0, \ldots, \tau_n, X_n \in \ell_0\}$$

Università di Pisa

# Motivation

Let

$$\widetilde{G}(n,x) = P\{\text{in } n \text{ steps and in time} \leq x \text{ the system goes from level } \ell_1 \text{ to level } \ell_0\}$$

Question: How fast does $\widetilde{G}(n,x)$ go to zero, as $n \longrightarrow \infty$?

Idea: Through Laplace-Stiltjes transforms reduce the MRP to an M/G/1-type Markov chain

# Motivation

The M/G/1-type Markov chain is defined by the transition probability matrix

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

Let $G_n = P\{\text{in } n \text{ steps we go from level } \ell_1 \text{to level } \ell_0\}$
Question: How fast does $G_n$ go to zero, as $n \to \infty$?

# Back to $G(w)$

Recall that $G(w)$ solves

$$X = w \sum_{n=-1}^{\infty} A_n X^{n+1}.$$

Known facts:

1. $G(w)$ is analytic $|w| < 1$, convergent for $|w| = 1$
2. $G(w) = \sum_{n=0}^{\infty} w^n G_n$
3. $G_n \geq 0$ for any $n \geq 0$, and $G_n$ is the sought probability.

# Back to $G(w)$

Recall that $G(w)$ solves

$$X = w \sum_{n=-1}^{\infty} A_n X^{n+1}.$$

Known facts:

1. $G(w)$ is analytic $|w| < 1$, convergent for $|w| = 1$
2. $G(w) = \sum_{n=0}^{\infty} w^n G_n$
3. $G_n \geq 0$ for any $n \geq 0$, and $G_n$ is the sought probability.

Problem: Can we find $\theta$ such that $\|G_n\| = O(\theta^n)$ as $n \to \infty$?

# Back to $G(w)$

Recall that $G(w)$ solves

$$X = w \sum_{n=-1}^{\infty} A_n X^{n+1}.$$

Known facts:

1. $G(w)$ is analytic $|w| < 1$, convergent for $|w| = 1$
2. $G(w) = \sum_{n=0}^{\infty} w^n G_n$
3. $G_n \geq 0$ for any $n \geq 0$, and $G_n$ is the sought probability.

Problem: Can we find $\theta$ such that $\|G_n\| = O(\theta^n)$ as $n \to \infty$?

Assumptions:

- $\sum_{i=-1}^{\infty} A_i$ is irreducible;
- the convergence radius of $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$ is $r_a > 1$.

# Idea

To apply to $G(w)$ the:

## Theorem (Cauchy estimate)

*Let $R$ be the convergence radius of the series $f(w) = \sum_{j=0}^{\infty} w^j f_j$. Then, for any $0 < r < R$ and for any $j \geq 0$ one has*

$$|f_j| \leq \frac{\mu(r)}{r^j}$$

*where $\mu(r) = \sup_{|w|=r} |f(w)|$.*

Problem: to determine the convergence radius of $G(w)$

# A scalar example: Poisson distribution

Let $\lambda > 0$ and

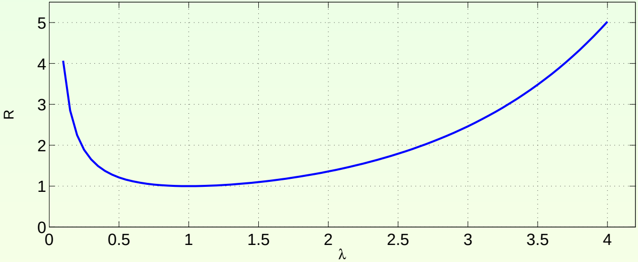
$$A_i = e^{-\lambda}\frac{\lambda^{i+1}}{(i+1)!}, \quad i \geq -1.$$

One has

$$G(w) = \sum_{n=1}^{\infty} w^n e^{-n(\lambda-1)}\lambda^{n-1}$$

The convergence radius is $R = (\lambda e^{-(\lambda-1)})^{-1}$

# Convergence radius $R$ in function of $\lambda$




How can we estimate $R$ without knowing $G(w)$?

# Scalar case

Let $A_i$ be scalar numbers, and $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$. We look for conditions on $w$ so that the equation $z = wA(z)$ has a solution in $(0, r_a)$.

The following properties hold:

- Since $A_i \geq 0$, the function $A(z)$ is convex and increasing in $(0, r_a)$;

# Scalar case

Let $A_i$ be scalar numbers, and $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$. We look for conditions on $w$ so that the equation $z = wA(z)$ has a solution in $(0, r_a)$.

The following properties hold:

- Since $A_i \geq 0$, the function $A(z)$ is convex and increasing in $(0, r_a)$;
- the function $f_w(z) = wA(z) - z$ has a minimum in $\sigma \in (0, r_a)$, where $\sigma$ solves the equation $wA'(\sigma) - 1 = 0$;

# Scalar case

Let $A_i$ be scalar numbers, and $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$. We look for conditions on $w$ so that the equation $z = wA(z)$ has a solution in $(0, r_a)$.

The following properties hold:

- Since $A_i \geq 0$, the function $A(z)$ is convex and increasing in $(0, r_a)$;
- the function $f_w(z) = wA(z) - z$ has a minimum in $\sigma \in (0, r_a)$, where $\sigma$ solves the equation $wA'(\sigma) - 1 = 0$;
- the equation $z = wA(z)$ has a solution in $(0, r_a)$ if $f_w(\sigma) \leq 0$.

Università di Pisa

# Scalar case

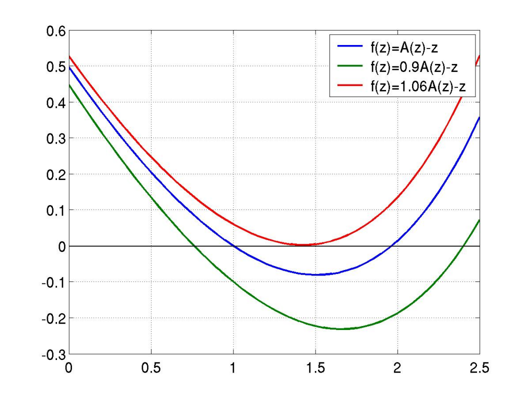
Let $A_i$ be scalar numbers, and $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$. We look for conditions on $w$ so that the equation $z = wA(z)$ has a solution in $(0, r_a)$.

The following properties hold:

- Since $A_i \geq 0$, the function $A(z)$ is convex and increasing in $(0, r_a)$;
- the function $f_w(z) = wA(z) - z$ has a minimum in $\sigma \in (0, r_a)$, where $\sigma$ solves the equation $wA'(\sigma) - 1 = 0$;
- the equation $z = wA(z)$ has a solution in $(0, r_a)$ if $f_w(\sigma) \leq 0$.

The convergence radius $R$ is the superior extremum of the $w \in (0, r_a)$ such that $z = wA(z)$ has a solution in $(0, r_a)$

# Scalar case




Here $R = 1.06$

# Convergence radius in the scalar case

The convergence radius of $G(w)$ is

$$R = \frac{1}{A'(\sigma)}$$

where $\sigma \in (0, r_a)$ solves the equation

$$A'(z)z = A(z)$$

# Convergence radius in the scalar case

The convergence radius of $G(w)$ is

$$R = \frac{1}{A'(\sigma)}$$

where $\sigma \in (0, r_a)$ solves the equation

$$A'(z)z = A(z)$$

Can we extend these properties to the block case?

# General case

Let $\mu_w$ and $\mathbf{u}_w$ be the Perron eigenvalue/eigenvector of $G(w)$.

# General case

Let $\mu_w$ and $\mathbf{u}_w$ be the Perron eigenvalue/eigenvector of $G(w)$. From

$$G(w)\mathbf{u}_w = w \sum_{i=-1}^{\infty} A_i G(w)^{i+1} \mathbf{u}_w$$

one has

$$\mu_w \mathbf{u}_w = w \left( \sum_{i=-1}^{\infty} A_i \mu_w^{i+1} \right) \mathbf{u}_w .$$

# General case

Let $\mu_w$ and $\mathbf{u}_w$ be the Perron eigenvalue/eigenvector of $G(w)$. From

$$G(w)\mathbf{u}_w = w\sum_{i=-1}^{\infty} A_i G(w)^{i+1}\mathbf{u}_w$$

one has

$$\mu_w \mathbf{u}_w = w\left(\sum_{i=-1}^{\infty} A_i \mu_w^{i+1}\right)\mathbf{u}_w.$$

Therefore, $\mu_w$ solves the *scalar equation*

$$z = w\rho(A(z)). \tag{1}$$

*Intuition: the convergence radius is the superior extremum of the* $w \in (0, r_a)$ *such that (1) has a solution*

# Some properties

Let $\theta(z)$ be the spectral radius of $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$.

The following properties hold [Gail,Hantler,Taylor 95]:

- $\theta(z)$ is a real analytic function in $(0, r_a)$;
- $\theta(z)$ is strictly increasing in $(0, r_a)$;
- the function $\log \theta(e^t)$ is convex and increasing in $(-\infty, \log r_a)$.

# Main result

## Theorem

*Let $\theta(z) = \rho(A(z))$. The following properties hold:*

1. *the equation*

$$\theta'(z)z = \theta(z)$$

*has a unique solution $\sigma$ in $(0, r_a)$;*

# Main result

**Theorem**

*Let $\theta(z) = \rho(A(z))$. The following properties hold:*

1. *the equation*

$$\theta'(z)z = \theta(z)$$

*has a unique solution $\sigma$ in $(0, r_a)$;*

2. *the equation $z = w\theta(z)$ has a solution in $(0, r_a)$ for any $0 < w \leq 1/\theta'(\sigma)$;*

# Main result

**Theorem**

*Let $\theta(z) = \rho(A(z))$. The following properties hold:*

1. *the equation*
$$\theta'(z)z = \theta(z)$$
*has a unique solution $\sigma$ in $(0, r_a)$;*
2. *the equation $z = w\theta(z)$ has a solution in $(0, r_a)$ for any $0 < w \leq 1/\theta'(\sigma)$;*
3. *the convergence radius of $G(w)$ is $R = 1/\theta'(\sigma) = \sigma/\theta(\sigma)$.*

# Fixed point iteration

Goal: Compute the solution $\sigma \in (0, r_a)$ of $z = \frac{\theta(z)}{\theta'(z)}$.

Set, say, $z_0 = 1$ and compute

$$z_{k+1} = \frac{\theta(z_k)}{\theta'(z_k)}, \quad k \geq 0.$$

# Fixed point iteration

Goal: Compute the solution $\sigma \in (0, r_a)$ of $z = \frac{\theta(z)}{\theta'(z)}$.

Set, say, $z_0 = 1$ and compute

$$z_{k+1} = \frac{\theta(z_k)}{\theta'(z_k)}, \quad k \geq 0.$$

At the $k$-th step we need:

- $\theta(z_k) = \rho(A(z_k))$ (easy to compute);

# Fixed point iteration

Goal: Compute the solution $\sigma \in (0, r_a)$ of $z = \frac{\theta(z)}{\theta'(z)}$.

Set, say, $z_0 = 1$ and compute

$$z_{k+1} = \frac{\theta(z_k)}{\theta'(z_k)}, \quad k \geq 0.$$

At the $k$-th step we need:

- $\theta(z_k) = \rho(A(z_k))$ (easy to compute);
- $\theta'(z_k)$?

# Computation of $\theta'(z)$

1. There exist vectors $\mathbf{u}(z)$ and $\mathbf{v}(z)$, analytic for $0 < z < r_a$, such that

$$\mathbf{u}(z)^T A(z) = \theta(z)\mathbf{u}(z)^T, \quad A(z)\mathbf{v}(z) = \theta(z)\mathbf{v}(z)$$

Università di Pisa

# Computation of $\theta'(z)$

1. There exist vectors $\mathbf{u}(z)$ and $\mathbf{v}(z)$, analytic for $0 < z < r_a$, such that

$$\mathbf{u}(z)^T A(z) = \theta(z)\mathbf{u}(z)^T, \quad A(z)\mathbf{v}(z) = \theta(z)\mathbf{v}(z)$$

2. Taking derivative in $A(z)\mathbf{v}(z) = \theta(z)\mathbf{v}(z)$ and multiplying the result on the left by $\mathbf{u}(z)^T$, yields

$$\theta'(z) = \frac{\mathbf{u}(z)^T A'(z)\mathbf{v}(z)}{\mathbf{u}(z)^T \mathbf{v}(z)}.$$

# Newton's method

Newton's method applied to the function $g(z) = \theta(z) - z\theta'(z)$, yields the sequence

$$z_{k+1} = z_k + \frac{\theta(z_k) - z_k\theta'(z_k)}{z_k\theta''(z_k)}, \quad k \geq 0,$$

starting, say, from $z_0 = 1$.

How compute $\theta''(z_k)$?

Università di Pisa

# Computation of $\theta''(z)$

Differentiate twice $A(z)\mathbf{v}(z) = \theta(z)\mathbf{v}(z)$, and multiply the result on the left by $\mathbf{u}(z)^T$.
We get

$$\theta''(z) = \frac{\mathbf{u}(z)^T \left(A''(z)\mathbf{v}(z) + 2A'(z)\mathbf{v}'(z) - 2\theta'(z)\mathbf{v}'(z)\right)}{\mathbf{u}(z)^T\mathbf{v}(z)},$$

where $\mathbf{v}'(z)$ solves the singular system

$$(A(z) - \theta(z)I)\mathbf{v}'(z) = -(A'(z) - \theta'(z))\mathbf{v}(z).$$

# Convergence

- We expect a local linear convergence for the fixed point iteration, and a local quadratic convergence for Newton's method.
- A detailed convergence analysis seems difficult to be performed.
- Experimentally, we observed that the convergence of the fixed point iteration obtained starting with $z_0 = 1$ is monotonic.

# Numerical experiments

Consider the case

$$A_{-1} = \begin{bmatrix} 0 & p \\ 0 & 0 \end{bmatrix}, \quad A_0 = \begin{bmatrix} 0.8-p & 0 \\ 1 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0.2 \\ 0 & 0 \end{bmatrix},$$

and $A_i = 0$ for $i > 1$, with $0 < p < 0.8$. This problem is positive recurrent if $p > 0.2$, and null recurrent if $p = 0.2$.

We implemented the algorithms in Matlab, and considered the cases obtained with $p = 0.1 \cdot k$, $k = 1, 2, \ldots, 7$.

The iterations have been halted if $|z_{k+1} - z_k| < 10^{-10}$.

Università di Pisa

# Numerical results

| $p$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 |
|---|---|---|---|---|---|---|---|
| Newton | 5 | 1 | 5 | 5 | 6 | 6 | 6 |
| F.P.I. | 81 | 1 | 52 | 47 | 43 | 39 | 36 |
| $R$ | 1.0135 | 1 | 1.0068 | 1.0222 | 1.0424 | 1.0657 | 1.0911 |

Table: Number of iterations and value of the radius $R$

# Further developments

- Convergence analysis of the algorithms
- Design of algorithms for computing an arbitrary number of matrices $G_n$
- Probabilistic interpretation of the formula $R = 1/\theta'(\sigma)$.