Markov jump linear systems

Pantelis Sopasakis

IMT institute for Advanced Studies Lucca

March 13, 2017

Outline

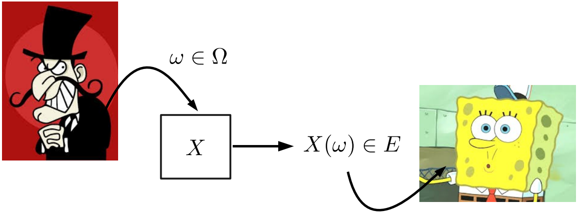
1. Quick probability theory brush-up
2. Markov jump linear systems
3. Expectation and covariance dynamics
4. Mean square stability (MSS)
5. Conditions for MSS
6. Almost sure convergence
7. Stabilisation of MJLS

I. Probability theory

1. Random variables
2. Conditional Probability
3. Expectation and Conditional Expectation
4. Filtration

Random variable

Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space and $(E, \mathcal{E})$ be a measurable space. A measurable function $X : \Omega \to E$ is called a **random variable**.



$\omega \in \Omega$

X

$X(\omega) \in E$

Expected value

The **expected value** $\mathbb{E}[X]$ of a random variable X is defined as

$$\mathbb{E}[X] = \int_\Omega X \mathrm{dP} = \int_\Omega X(\omega)\, \mathrm{P}(\mathrm{d}\omega).$$

Conditional probability

On a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ and for $A, B \in \mathcal{F}$ with $\mathrm{P}(B) > 0$ we define the **conditional probability** of A given B as

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

Let X be a random variable; then we define the conditional probability $\mathrm{P}(A \mid X)$ to be a random variable with

$$\mathrm{P}(A \mid X)(\omega) = \mathrm{P}(A \mid X = X(\omega)).$$

Conditional expectation

Given a random variables X on a prob. space $(\Omega, \mathcal{F}, \mathrm{P})$ and a $H \in \mathcal{F}$, we define the **conditional expecation** $\mathbb{E}[X \mid H]$ to be

$$\mathbb{E}[X \mid H] = \int_{\Omega} X(\omega)\, \mathrm{P}(\mathrm{d}\omega \mid H).$$

Conditional expectation

Given a random variables X on a prob. space $(\Omega, \mathcal{F}, \mathrm{P})$ and a sub-σ-algebra $\mathcal{H} \subseteq \mathcal{F}$ we call **conditional expecation** of X given $\mathcal{H}$ as a measurable function $\mathbb{E}[X \mid \mathcal{H}]$ such that

$$\int_H \mathbb{E}[X \mid \mathcal{H}] \mathrm{dP} = \int_H X \mathrm{dP}, \ \forall H \in \mathcal{H}.$$

The **conditional expectation** of X wrt another random variable Y is the random variable

$$\mathbb{E}[X \mid Y = y] = \int_\Omega X(\omega) \, \mathrm{P}(\mathrm{d}\omega \mid Y = y) = \mathbb{E}[X \mid \sigma(Y)].$$

Stochastic process

A (discrete) **stochastic process** is a sequence of random variables from a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ to a measurable space $(E, \mathcal{E})$, that is $\{X_k\}_{k\in\mathbb{N}}$.

Filtration

Formally, a **filtration** over a prob. space $(\Omega, \mathcal{F}, \mathrm{P})$ is a sequence of sub-σ-algebras of $\mathcal{F}$, $\{F_k\}_{k\in\mathbb{N}}$ so that

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}.$$

Filtration

Formally, a **filtration** over a prob. space $(\Omega, \mathcal{F}, \mathrm{P})$ is a sequence of sub-σ-algebras of $\mathcal{F}$, $\{F_k\}_{k\in\mathbb{N}}$ so that

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}.$$

- The space $(\Omega, \mathcal{F}, \{\mathcal{F}_k\}_k, \mathrm{P})$ is called a **filtered** prob. space.

Filtration

Formally, a **filtration** over a prob. space $(\Omega, \mathcal{F}, \mathrm{P})$ is a sequence of sub-σ-algebras of $\mathcal{F}$, $\{F_k\}_{k \in \mathbb{N}}$ so that

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}.$$

- The space $(\Omega, \mathcal{F}, \{\mathcal{F}_k\}_k, \mathrm{P})$ is called a **filtered** prob. space.
- A filtration $\{F_k\}_{k \in \mathbb{N}}$ represents the **flow of information** as the experiment goes on in time,

Filtration

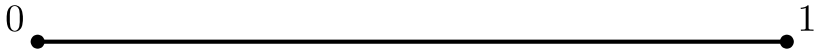
Formally, a **filtration** over a prob. space $(\Omega, \mathcal{F}, \mathrm{P})$ is a sequence of sub-σ-algebras of $\mathcal{F}$, $\{F_k\}_{k \in \mathbb{N}}$ so that

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}.$$

- The space $(\Omega, \mathcal{F}, \{\mathcal{F}_k\}_k, \mathrm{P})$ is called a **filtered** prob. space.
- A filtration $\{F_k\}_{k \in \mathbb{N}}$ represents the **flow of information** as the experiment goes on in time,
- $\mathcal{F}_k$ encodes the **available information** up to time k,

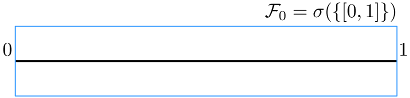
Filtration – example

Take $\Omega = [0, 1]$ and $\mathcal{F} = \sigma(\Omega)$.



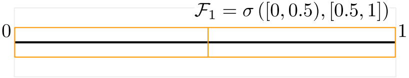
Filtration – example

Take $\mathcal{F}_0 = \{\varnothing, [0,1]\}$ (trivially $\mathrm{P}(\varnothing) = 0$ and $\mathrm{P}([0,1]) = 1$)



Filtration – example

Take $\mathcal{F}_1 \supseteq \mathcal{F}_0$ as follows



Filtration – example

Now construct a $\mathcal{F}_2 \supseteq \mathcal{F}_1 \supseteq \mathcal{F}_0$ as follows

$\mathcal{F}_1$

$$\mathcal{F}_2 = \sigma\left(\boxed{[0,1/6),[1/6,1/3),[1/3,1/2)},\boxed{[1/2,1]}\right)$$

Adapted processes

We say that a stochastic process $\{X_k\}_{k\in\mathbb{N}}$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_k\}_k, \mathrm{P})$ is **adapted** to the filtration $\{\mathcal{F}_k\}_k$ if every X_k is $\mathcal{F}_k$-measurable. Equivalently:

$$\mathbb{E}[X_k \mid \mathcal{F}_k] = X_k.$$

Natural filtration

Given $\{X_k\}_k$ we can construct a filtration so that the given random process is adapted to it. This is called the **natural filtration** and it is $\mathcal{F}_k = \sigma(\cup\{\sigma(X_j)\}_{j\leq k})$.

Useful properties

- **Markov's inequality.** If X is a positive, integrable random variable,

$$\mathrm{P}[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}.$$

- **Chebychev's inequality.** If X a finite-variance random variable

$$\mathrm{P}[|X - \mathbb{E}[X]| \geq \alpha] \leq \frac{\mathrm{Var}[X]}{\alpha^2}.$$

- **Jensen's inequality.** For φ convex and X integrable r.v.

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)].$$

The Borel-Cantelli lemma

Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space and $\{E_i\}_i$ a sequence of events. We denote by $\limsup_i E_i$ the following

$$\limsup_{i\to\infty} E_i = \bigcap_{k\geq 1} \bigcup_{j\geq k} E_j$$

Notice that

- this a set-theoretic entity and does assume/require a topology
- it is different from the limit-superior used to define the Painleve-Kuratowski limit
- $x \in \limsup_{i\to\infty} E_i$ iff x is in every $\cup_{j\geq k} E_j$
- $\limsup_{i\to\infty} E_i \in \mathcal{F}$

The Borel-Cantelli lemma

Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space and $\{E_i\}_i$ a sequence of events. If

$$\sum_{i=1}^{\infty} \mathrm{P}[E_i] < \infty,$$

then

$$\mathrm{P}[\limsup_i E_i] = 0.$$

Further reading

Excellent references:

1. S.R.S.Varadhan, *Probability theory*, Lecture notes, NY Univ., 2000. Available online at https://goo.gl/ygbqa3.
2. Erhan Çınlar, *Probability and stochastics*, Springer, 2011.

Summary of first section

We revised the notions

1. Probability and measure spaces, probability, conditional probability
2. Expectation and Conditional expectation
3. Stochastic process
4. Filtration and adaptation to a filtration

II. Markov Jump Linear Systems

1. Definition of MJLS
2. Applications
3. MJLS quirks

Markov processes

A *Markov process* is a stochastic process[1] so that

$$\mathrm{P}\left[w(k)=\omega|w(0),w(1),\ldots,w(k-1)\right]=\mathrm{P}\left[w(k)=\omega|w(k-1)\right].$$

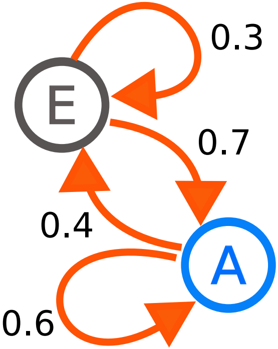
We assume our Markov processes are:

1. finite: $w(k)$ is finite valued $w(k)\in\mathcal{N}=\{1,\ldots,N\}$
2. time-homogeneous: $\mathrm{P}\left[w(k)=j|w(k-1)=i\right]=p_{ij}$.

---

[1] We call *stochastic process a sequence of random variables* $w(k)$ *over a probability space* $(\Omega,\mathcal{F},\mathrm{P})$.

Finite Markov processes



The transition matrix is

$$P = \begin{bmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{bmatrix}$$

with

$$\mathcal{N} = \{1, 2\}$$

Markov jump linear systems

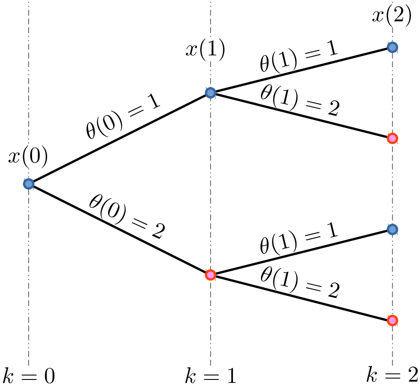
A (autonomous) Markov jump linear system is a dynamical system of the form

$$x(k+1) = \Gamma_{\theta(k)} x(k),$$

where $\{\theta(k)\}_k$ is a finite time-homogeneous Markov process, $x(0) = x_0$ and $\theta(0) = \theta_0$ follows an initial distribution with $\mathrm{P}[\theta(0) = \omega_i] = v_i$ for $i = 1, \ldots, N$.

Markov jump linear systems

Notice that $\{x(k)\}_k$ is not a Markov process, but $\{(x(k), \theta(k))\}_k$ is.



MJLS applications

Examples

1. Power systems [Li *et al.* 2007; Ugrinovskii, 2005]
2. Satellite control [Meskin & Khorasani,2009]
3. Flight control [Gray, González & Doğan, 2000]
4. Air traffic mgmt [Zhou, 2011]
5. Solar thermal receiver [Sworder & Rogers, 1983]
6. Cell growth [Horowitz *et al.*, 2010]
7. Samuelson's macroeconomic model [Costa *et al.*, 1999]
8. Printer over network with random time delays [Patrinos *et al.*, 2011]

MJLS quirks

Assume that **some or all Γ_i are unstable**. How will the system trajectories look like? Example:

$$\Gamma_1 = \begin{bmatrix} 2 & -1 \\ 0.1 & 0.1 \end{bmatrix}, \Gamma_2 = \begin{bmatrix} 0.2 & 1 \\ -0.1 & 2 \end{bmatrix},$$

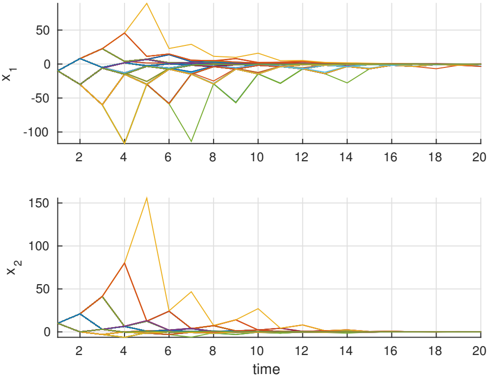
and

$$P = \begin{bmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{bmatrix}.$$

Take $x(0) = (1, 1)'$ and $\mathrm{P}\left[w(0) = 1\right] = 0.5$.

Unstable + Unstable = Stable?!

Stable trajectories of a MJLS with unstable subsystems.



MJLS quirks continued

Assume that Γ_i are strictly Hurwitz for all $i = 1, \ldots, N$ (**all stable**) – how will the MJLS Trajectories look like? Will they all converge to zero?
Example:

$$\Gamma_1 = \begin{bmatrix} -0.5 & 2 \\ -0.5 & 0.5 \end{bmatrix}, \Gamma_2 = \begin{bmatrix} 0.5 & 0 \\ 1 & 0 \end{bmatrix},$$

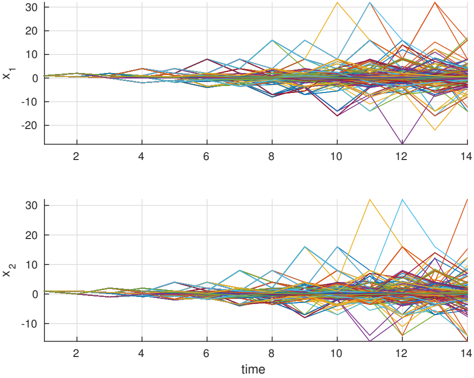
for which $\rho(\Gamma_1) = 0.866$ and $\rho(\Gamma_2) = 0.5$, and

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

Take $x(0) = (1, 1)'$ and $\mathrm{P}\left[w(0) = 1\right] = 0.5$.

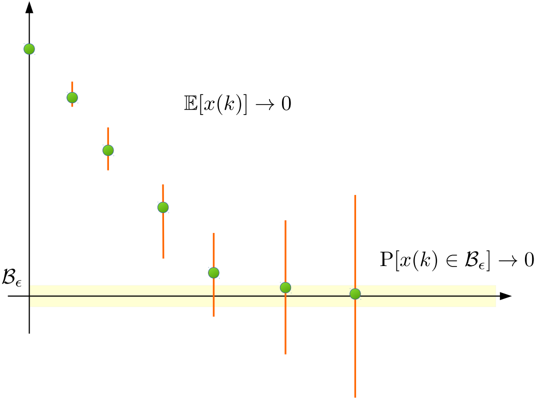
Stable + Stable = Unstable?!

While the expected value $\mathbb{E}[\|x(k)\|]$ converges to 0, the covariance diverges to infinity!



Is $\mathbb{E}[x(k)] \to 0$ enough?

"Averages are funny things. Ask the statistician who drowned in a lake of average depth 3 centimetres." ∼ J. Barrow, 100 things you didn't know you didn't know.



End of second section

1. We learnt what a Markov process is
2. and the class of Markov jump linear systems
3. We then listed application of Markov jump systems
4. We realised there's a lot to learn regarding MJLS...

III. MJLS dynamics

Questions

1. How do we model the evolution of a MJLS?
2. What definition(s) of stability apply to MJLS?
3. How do we control such a system?

Some definitions

For a set $F \in \mathcal{F}$ we define its characteristic function $\chi_F : \Omega \to \{0, 1\}$ as

$$\chi_F(\omega) := \begin{cases} 1, & \text{if } \omega \in F \\ 0, & \text{otherwise} \end{cases}$$

Then, notice that

$$\mathbb{E}[x(k)] = \sum_{i=1}^{N} \mathbb{E}[x(k)\chi_{\{\theta(k)=i\}}]$$

and

$$\mathbb{E}[x(k)x(k)'] = \sum_{i=1}^{N} \mathbb{E}[x(k)x(k)'\chi_{\{\theta(k)=i\}}]$$

We are going to describe the dynamics of $\mathbb{E}[x(k)]$ and $\mathbb{E}[x(k)x(k)']$.

Some useful Hilbert spaces

Spaces $\mathbb{R}^n$ and $\mathbb{R}^{n \times m}$ are Hibert spaces with inner products

$$\langle x, y \rangle := x'y, \text{and } \langle X, Y \rangle := \operatorname{tr}(X'Y)$$

We define the Hilbert space $\mathbb{H}^{n,m}$ of N-tuples of m-by-n matrices, i.e., $H = (H_1, \ldots, H_N)$ with $H_i \in \mathbb{R}^{m \times n}$ with the following inner product:

$$\langle H, V \rangle := \sum_{i=1}^{N} \langle H_i, V_i \rangle$$

On $\mathbb{H}^{n,m}$ we define the norms

$$\|H\|_1 := \sum_{i=1}^{N} \|H_i\|_2, \text{and } \|H\|_2 := \left(\sum_{i=1}^{N} \underbrace{\langle V_i, V_i \rangle}_{\|V_i\|_{\mathrm{f}}} \right)^{\frac{1}{2}}$$

Some useful Hilbert spaces

We denote by $C^n = L^2(\omega, \mathcal{F}, \mathrm{P})$ the space of $\mathbb{R}^n$-valued random variables y with inner product $\langle y, z \rangle = \mathbb{E}[y'z]$ so that $\|y\| < \infty$.

We call $\ell_2(C^n)$ the space of sequences $s = \{s(k)\}_k$ with $s(k) \in C^n$ and[2]

$$\|s\|_{\ell_2(C^n)} := \sum_{k \in \mathbb{N}} \|s(k)\|_{C^n}^2 < \infty.$$

---

[2] Note that $\|s(k)\|_{C^n}^2 = \langle s(k), s(k) \rangle_{C^n} = \mathbb{E}[\|s(k)\|^2]$.

Some useful Hilbert spaces

Now let $(\Omega, \mathcal{F}, \{\mathcal{F}_k\}_{k\in\mathbb{N}}, \mathrm{P})$ be the aforementioned probability space with a *filtration* $\{\mathcal{F}_k\}_k$. We define the space $\mathcal{C}^n \subseteq \ell_2(C^n)$ if for every random process $r \in \mathcal{C}^n$, $r(k)$ is $\mathcal{F}_k$-measurable, i.e., r is *adapted* to the filtration $\{\mathcal{F}_k\}_k$.

MJLS dynamics

We are now going to describe the dynamics of $\mathbb{E}[x(k)]$ and $\mathbb{E}[x(k)x(k)']$. Define

$$q_i(k) = \mathbb{E}\left[x(k)\chi_{\{\theta(k)=i\}}\right]$$

And defining $q(k) := (q_1(k), \ldots, q_2(k)) \in \mathbb{H}^{1,n}$ we have

$$\mu(k) := \mathbb{E}[x(k)] = \sum_{i=1}^{N} q_i(k).$$

MJLS dynamics

We are now going to describe the dynamics of $\mathbb{E}[x(k)]$ and $\mathbb{E}[x(k)x(k)']$. Define

$$q_i(k) = \mathbb{E}\left[x(k)\chi_{\{\theta(k)=i\}}\right]$$

And defining $q(k) := (q_1(k), \dots, q_2(k)) \in \mathbb{H}^{1,n}$ we have

$$\mu(k) := \mathbb{E}[x(k)] = \sum_{i=1}^{N} q_i(k).$$

Define also

$$Q_i(k) := \mathbb{E}\left[x(k)x(k)'\chi_{\{\theta(k)=i\}}\right],$$

and let $Q(k) := (Q_1(k), \dots, Q_N(k)) \in \mathbb{H}^{n,n}$, then

$$\Sigma(k) := \mathbb{E}\left[x(k)x(k)'\right] = \sum_{i=1}^{N} Q_i(k).$$

MJLS dynamics

The mean and covariance dynamics are described by

$$q_j(k+1) = \sum_{i=1}^{N} p_{ij} \Gamma_i q_i(k)$$

and

$$Q_j(k+1) = \sum_{i=1}^{N} p_{ij} \Gamma_i Q_i(k) \Gamma_i'.$$

In what follows we will define linear operators $\mathcal{B} : q(k) \mapsto q(k+1)$ and $\mathcal{T} : Q(k) \mapsto Q(k+1)$.

A bound on $\|x(k)\|^2$

We can show that[3]

$$\|x(k)\|^2 \leq n\|Q(k)\|_1,$$

where notice that the left hand side norm is the norm of C^n, whereas the right hand side is the 1-norm in $\mathbb{H}^{n,n}$. What happens as $Q(k) \to 0$?

---

[3] In $C^n = L^2(\Omega, \mathcal{F}, \mathrm{P}, \mathrm{I\!R}^n)$ we defined the inner product $\langle x, y\rangle = \mathbb{E}[x'y]$, so the induced squared norm is $\|x\|^2 = \langle x, x\rangle = \mathbb{E}[x'x] = \mathbb{E}[\|x\|^2]$.

Operators $\mathcal{T}$ and $\mathcal{L} = \mathcal{T}^*$

We have $Q(k+1) = \mathcal{T}[Q(k)]$ where $\mathcal{T} : \mathbb{H}^n \to \mathbb{H}^n$ is the linear operator $\mathcal{T}[Q] := (\mathcal{T}_1[Q], \dots, \mathcal{T}_N[Q])$, with

$$\mathcal{T}_j[Q] := \sum_{i=1}^{N} p_{ij} \Gamma_i Q_i \Gamma_i'.$$

The adjoint of this operator is an operator $\mathcal{L}[Q] := (\mathcal{L}_1[Q], \dots, \mathcal{L}_N[Q])$, with

$$\mathcal{L}_i[Q] := \Gamma_i \mathcal{E}_i[Q] \Gamma_i,$$

where $\mathcal{E}_i[Q] = \sum_{i=j}^{N} p_{ij} Q_j$[4].

---

[4] Use the definition, i.e., show that $\langle \mathcal{T}[Q], S \rangle = \langle Q, \mathcal{L}[S] \rangle$ for all $Q, S \in \mathbb{H}^n$. It is left as an exercise to prove that $\mathcal{T}^* = \mathcal{L}$.

Operator $\mathcal{B}$ as a matrix

We have $q(k+1) = \mathcal{B}q(k)$, where $\mathcal{B} : \mathbb{R}^{Nn} \to \mathbb{R}^{Nn}$ is a linear operator[5] and is given by

$$\mathcal{B}[q] = (P' \otimes I_n)\,\mathrm{blkdiag}(\Gamma_1, \ldots, \Gamma_N) \cdot q.$$

If $\rho(\mathcal{B}) < 1$ then $q(k) \to 0$ and $\mu(k) \to 0$, **however**, this does not imply $Q(k) \to 0$ or $\Sigma(k) \to 0$.

---

[5] Spaces $\mathbb{R}^{Nn}$ and $\mathbb{H}^{1,n}$ are homeomorphic (can be thought of as the same space).

Operator $\mathcal{T}$ as a matrix

For a matrix $A \in \mathbb{R}^{m \times n}$, $A = \left[\begin{smallmatrix} A_1 & \cdots & A_n \end{smallmatrix} \right]$ we define

$$\mathrm{vec}(A) := \left[\begin{smallmatrix} A_1' & \cdots & A_n' \end{smallmatrix} \right]' \in \mathbb{R}^{mn}.$$

For $H \in \mathbb{H}^{n,m}$ with $H = (H_1, \ldots, H_N)$, $H_i \in \mathbb{R}^{m \times n}$

$$\mathrm{vec}(H) := \left[\begin{smallmatrix} \mathrm{vec}(H_1)' & \cdots & \mathrm{vec}(H_n)' \end{smallmatrix} \right]' \in \mathbb{R}^{Nmn}.$$

We are looking for a matrix $T \in \mathbb{R}^{Nn^2 \times Nn^2}$ so that

$$\mathrm{vec}(\mathcal{T}[Q]) = T \, \mathrm{vec}(Q)$$

We may verify that

$$T = (P' \otimes I_{n^2}) \, \mathrm{blkdiag}(\{\Gamma_i' \otimes \Gamma_i\}_i).$$

If $\mathcal{T}$ is stable, then $\mathcal{B}$ is stable

We've seen that $q(k) \to 0$ does not imply $Q(k) \to 0$. However, the converse holds true: If $\mathcal{T}$ is stable, then $\mathcal{B}$ is stable too!

$$\rho(\mathcal{T}) < 1 \Rightarrow \rho(\mathcal{B}) < 1.$$

End of third section

1. We described the dynamics of $\mathbb{E}[x(k)]$ and $\mathbb{E}[x(k)x(k)']$
2. We introduced the Hilbert spaces $\mathbb{H}^{n,m}$, C^n, $\ell_2(C^n)$ and $\mathcal{C}^n$
3. We defined $q(k)$ and $Q(k)$ to study the dynamics of $\mathbb{E}[x(k)]$ and $\mathbb{E}[x(k)x(k)']$
4. We introduced the operators $\mathcal{B}$ and $\mathcal{T}$ using which we can study the stability properties of the dynamics of $q(k)$ and $Q(k)$.

IV. Mean square stability

Mean square stability

The MJLS $x(k+1) = \Gamma_{\theta(k)} x(k)$ – where $\{\theta(k)\}_{k \in \mathbb{N}}$ is a Markov process – is **mean square stable** (MSS) if there is a μ and a Σ so that for any initial state $x(0) = x_0$ and initial distribution of $\theta(0)$ it is

1. $\mu(k) \to \mu$,
2. $\Sigma(k) \to \Sigma$.

In our case, $\mu = 0$ and $\Sigma = 0$ and #2 implies #1.

Why MSS?

1. It's easy to test for
2. It implies stability of the expected state dynamics
3. It implies convergence in probability, i.e., for all $\epsilon > 0$,

   $$\mathrm{P}[\|x(k)\| \geq \epsilon] \to 0,$$

   and further
4. It implies almost sure convergence, i.e.,

   $$\mathrm{P}[\|x(k)\| \to 0] = 1.$$

MSS $\Rightarrow$ Convergence of $\mathbb{E}\left[x(k)\right]$

MSS $\Rightarrow$ Convergence of $\mathbb{E}\left[x(k)\right]$.

Already clear from the fact that $\rho(\mathcal{T}) < 1 \Rightarrow \rho(\mathcal{B}) < 1.$ Easy way to show it by Jensen's inequality:

$$0 \leq \mathbb{E}\left[\|x(k)\|\right] \leq \mathbb{E}\left[\|x(k)\|^2\right]^{\frac{1}{2}}.$$

BUT: Convergence of $\mathbb{E}\left[x(k)\right] \not\Rightarrow$ MSS (Exercise).

MSS $\Leftrightarrow \mathcal{T}$ is stable

If we know $\mathcal{T}$ (or T) we can tell whether our MJLS is stable.

If $\mathcal{T}$ is stable $\Leftrightarrow$ MJLS is MSS.

MSS $\Leftrightarrow \mathcal{T}$ is stable

If we know $\mathcal{T}$ (or T) we can tell whether our MJLS is stable.

If $\mathcal{T}$ is stable $\Leftrightarrow$ MJLS is MSS.

Proof.

Indeed, we have $Q(k+1) = \mathcal{T}Q(k)$ and $Q(k) = \mathcal{T}^k Q(0)$, so

$$\Sigma(k) = \sum_{i=1}^{N} Q_i(k) = \sum_{i=0}^{N} \mathcal{T}_i^k Q_i(0)$$

Since $\mathcal{T}$ is stable, $\mathcal{T}^k[Q] \to 0$, thus $\Sigma(k) \to 0$. □

The converse is very easy to show.

MSS $\Leftrightarrow$ MSES

MSS systems are *mean square exponentially stable*, i.e., there is a $\beta \geq 1$ and $\zeta \in (0, 1)$ so that

$$\mathbb{E}\left[\|x(k)\|^2\right] \leq \beta\zeta^k \|x_0\|^2, \ \forall k \in \mathbb{N}$$

Proof.

To prove that an MSS system is also MSES we need the following result from analysis: Let $\mathcal{T}$ be a linear operator with $\rho(\mathcal{T}) < 1$. Then there are $\beta \geq 1$ and $\zeta \in (0, 1)$ so that

$$\|\mathcal{T}^k\| \leq \beta\zeta^k,$$

where on the LHS we may have any operator norm. The rest is left as an exercise. □

Unstable + Unstable = Stable?

Two unstable modes:

$$\Gamma_1 = \begin{bmatrix} 2 & -1 \\ 0.1 & 0.1 \end{bmatrix}, \Gamma_2 = \begin{bmatrix} 0.2 & 1 \\ -0.1 & 2 \end{bmatrix},$$

with $\rho(\Gamma_1) = 1.9458$ and $\rho(\Gamma_2) = 1.9426$ and

$$P = \begin{bmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{bmatrix}.$$

but $\rho(\mathcal{T}) = \rho(T) = 0.6798 < 1$.

Stable + Stable = Unstable?

Two stable modes:

$$\Gamma_1 = \begin{bmatrix} -0.5 & 2 \\ -0.5 & 0.5 \end{bmatrix}, \Gamma_2 = \begin{bmatrix} 0.5 & 0 \\ 1 & 0 \end{bmatrix},$$

for which $\rho(\Gamma_1) = 0.866$ and $\rho(\Gamma_2) = 0.5$, and

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

but $\rho(\mathcal{T}) = \rho(T) = 1.0165 > 1$.

Stochastic stability

A MJLS is said to be stochastically stable (SS) if for all $x(0)$ and initial distributions for $\theta(0)$ it is

$$\sum_{k=0}^{\infty} \mathbb{E}\left[\|x(k)\|^2\right] < \infty.$$

This is equivalent to the random process $x = \{x(k)\}_k$ being in $\mathcal{C}^n$.

MSS $\Leftrightarrow$ SS

A MJLS is MSS iff it is SS[6].

Proof.

Since $\mathbb{E}\left[\|x(k)\|^2\right]$ is a sequence of ℓ^2, it converges $\mathbb{E}\left[\|x(k)\|^2\right] \to 0$. But we have

$$\begin{aligned} 0 \leq \mathbb{E}\left[\|x(k)\|^2\right] &= \mathbb{E}\left[\operatorname{tr}(x(k)x(k)')\right] \\ &= \operatorname{tr}\mathbb{E}\left[(x(k)x(k)')\right] \\ &= \operatorname{tr} Q(k) \\ &= \Sigma(k), \end{aligned}$$

so $\Sigma(k) \to 0$ for any $x(0)$ and $\theta(0)$. □

---

[6] This is not true for infinite Markov jump linear systems.

Positive (semi)definiteness

An $E \in \mathbb{R}^{n \times n}$ is *positive semidefinite* (PSD) if $x'Ax \geq 0$, $\forall x \in \mathbb{R}^n$; we denote $A \geq 0$. E is *positive definite* (PD) if it is PSD and $x'Ax = 0$ only when $x = 0$. Space of PSD (PD) matrices: S^n_+ (S^n_{++}).

It is $A < B$ whenever $B - A > 0$.

$\mathbb{H}^n_+$: the subset of $\mathbb{H}^n$ so that $H \in \mathbb{H}^n_+$ if $H_i \in S^n_+$ for all i.

Lyapunov-like stability conditions

1. A MJLS is MSS iff for every $S \in \mathbb{H}^{n+}$, $S > 0$, there exists a unique $V \in \mathbb{H}^{n+}$, $V > 0$ such that
$$V - \mathcal{T}(V) = S.$$
2. A MJLS is MSS iff for some $V \in \mathbb{H}^{n+}$, $V > 0$ it is
$$\mathcal{T}(V) < V.$$

Notice that $V - \mathcal{T}(V) = S$ is equivalent to $(\mathrm{id} - \mathcal{T})V = S$, which using von Neuman's identity yields $V = \sum_{k=0}^{\infty} \mathcal{T}^k [S]$.

Lyapunov-like stability conditions

A MJLS is MSS iff for every $S \in \mathbb{H}^{n+}$, $S > 0$, there exists a unique $V \in \mathbb{H}^{n+}$, $V > 0$ such that $V - \mathcal{T}(V) = S$.

Proof.

Sketch: we define the following system

$$Y(k+1) = \mathcal{L}(Y(k));\ Y(0) \in \mathbb{H}^{n}_{+}.$$

We need to prove that this is stable using the following Lyapunov function:

$$\Phi(Y) = \langle V, Y \rangle$$

Then $\mathcal{T} = \mathcal{L}^*$ will also be stable. □

MSS conditions using $\mathcal{L}$

The first condition can also be written as

$$V = \mathcal{L}(V) + S.$$

and now we have

$$V = \sum_{k=0}^{\infty} \mathcal{L}^k \left[S\right]$$

MSS conditions using $\mathcal{V}$ and $\mathcal{J}$

We have defined the operators $\mathcal{T}$ and $\mathcal{L}$ as

$$\mathcal{T}_j[Q] := \sum_{i=1}^{N} p_{ij}\Gamma_i Q_i \Gamma_i', \text{ and } \mathcal{L}_i[Q] := \Gamma_i \mathcal{E}_i[Q]\Gamma_i.$$

We now define the (**simpler**) operators $\mathcal{V}$ and $\mathcal{J}$[7] as

$$\mathcal{V}_j[Q] := \sum_{i=1}^{N} p_{ij}\Gamma_j Q_i \Gamma_j', \text{ and } J_i[Q] := \sum_{j=1}^{N} \Gamma_j' Q_j \Gamma_j.$$

Then, the MSS conditions hold also if we replace $\mathcal{T}$ with either $\mathcal{V}$ or $\mathcal{J}$.

---

[7] It is $\mathcal{V}^* = \mathcal{J}$.

Outlook...

These stability conditions are useful:

1. For **controller design** using linear matrix inequalities (LMIs)
2. To devise **stabilizability** and **detectability** tests
3. For MJLS with additive uncertainty
4. Similar conditions apply to **nonlinear** Markovian switching systems

Understanding MSS

Let $\{x(k)\}_k$ be the response of a MSS MJLS – all $x(k)$ are random variables, i.e., $x(k)(\omega)$. Let $\{\theta(k)\}_k$ be the corresponding sequence of random parameters. Can there be a realisation $\{\theta_k^\star\}$ so that the corresponding state trajectory

$$x(k; x(0), \theta(0), \ldots, \theta(k-1)) \to \infty \text{ as } k \to \infty \ ?$$

Understanding MSS

Let $\{x(k)\}_k$ be the response of a MSS MJLS – all $x(k)$ are random variables, i.e., $x(k)(\omega)$. Let $\{\theta(k)\}_k$ be the corresponding sequence of random parameters. Can there be a realisation $\{\theta_k^\star\}$ so that the corresponding state trajectory

$$x(k; x(0), \theta(0), \ldots, \theta(k-1)) \to \infty \text{ as } k \to \infty \text{ ?}$$

Short answer: **Yes!** (But the probability of running into one is zero...)

End of fourth section

Summary:

1. MSS is a “more meaningful” (practical) stability notion than MS
2. MSS: Both $\mu(k)$ and $\Sigma(k)$ converge
3. MSS $\Leftrightarrow$ $\mathcal{T}$ is stable
4. MSS $\Leftrightarrow$ MSES
5. MSS $\Leftrightarrow$ SS
6. MSS $\Leftrightarrow$ $\mathcal{T}V < V$, and $V - \mathcal{T}V = S \in \mathbb{H}^{n+}$

V. Almost sure convergence

Zero probability for non converging $x(k)$

MSS $\Rightarrow$ The probability that we find a realisation of $x(k)$ that does not converge is zero.

When is this probability zero?

If this probability is zero, do we have MSS?

Almost sure convergence

Under what conditions do we have $x(k) \to 0$ with probability one?

We say that $\{x(k)\}_k$ converges **almost surely** (or with probability 1) to 0 if[8]

$$\mathrm{P}(\lim_k \|x(k)\| = 0) = 1.$$

For a MJLS, MSS implies almost sure convergence – the converse is not true. Almost sure convergence is weaker than MSS[9].

---

[8] This probability is over the space of all random processes $\{x(k)\}_k$ which satisfy the MJLS dynamics.

[9] This type of convergence is not induced by some metric and does not depend on any topology.

Almost sure convergence

MSS $\Rightarrow$ Almost sure convergence

Proof.

We have $\mathbb{E}\left[\|x(k)\|^2\right] \leq \beta\zeta^k\|x_0\|^2$ for all $k \in \mathbb{N}$, so $\sum_{k=0}^{\infty} \mathbb{E}\left[\|x(k)\|^2\right]$ is finite and the rest follows from the Borel-Cantelli[10] lemma using the Chebychev-Markov inequality (Exercise). □

---

[10] From the Borel-Cantelli lemma we have that $z(k) \to 0$ almost surely (w.p.1) whenever $\sum_{k\in\mathbb{N}} \mathrm{P}\left[\|z(k)\| \geq \epsilon\right] < \infty$ for all $\epsilon > 0$.

ASC $\not\Rightarrow$ MSS

ASC does not imply MSS! Example with $x \in \mathbb{R}$, $N = 2$:

$$\Gamma_1 = 2.5, \text{and } \Gamma_2 = 0.1,$$

and

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

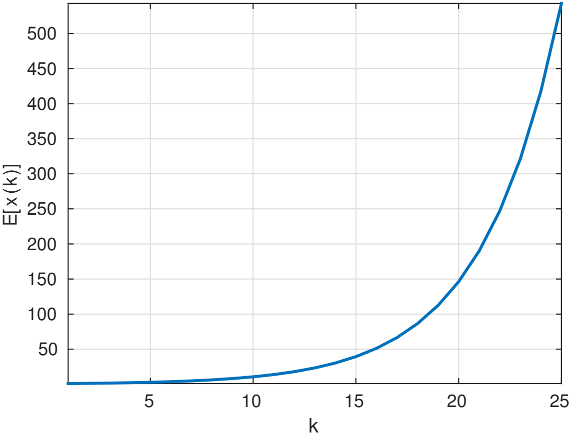
Then this is not MSS, but it converges almost surely.[11]

Interestingly, $\rho(\mathcal{T}) = 3.13$ (**not MSS**), but also $\rho(\mathcal{B}) = 1.3$ (**not MS**)!

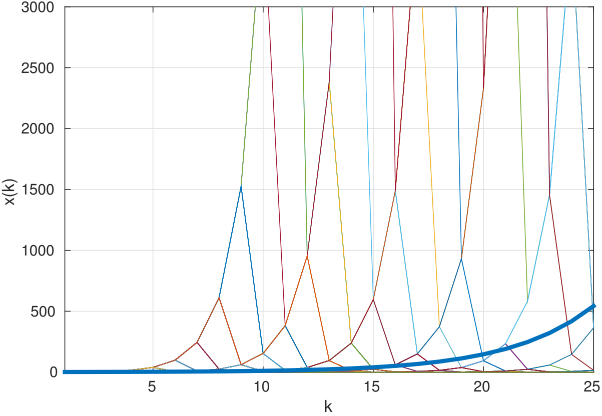
---

[11] See: Fang *et al.* 1994.

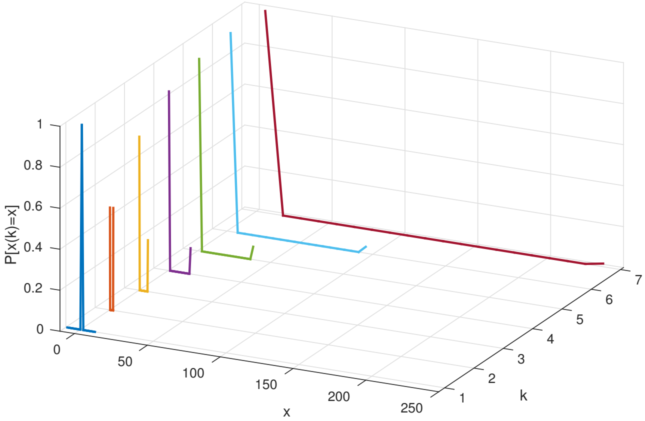
ASC $\not\Rightarrow$ MSS



Understanding ASC



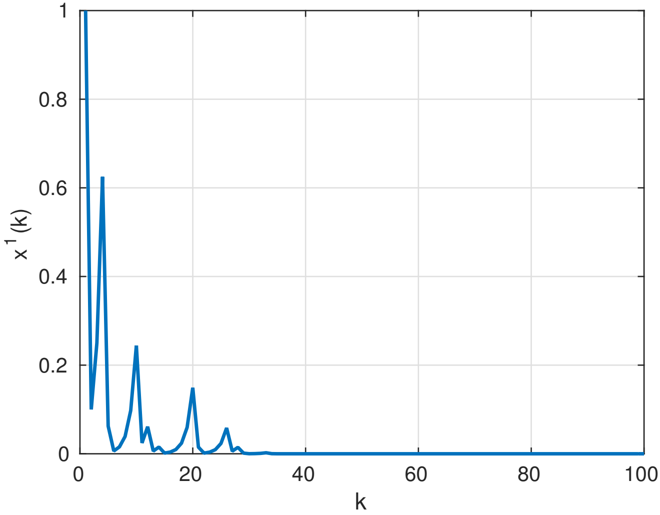
Understanding ASC



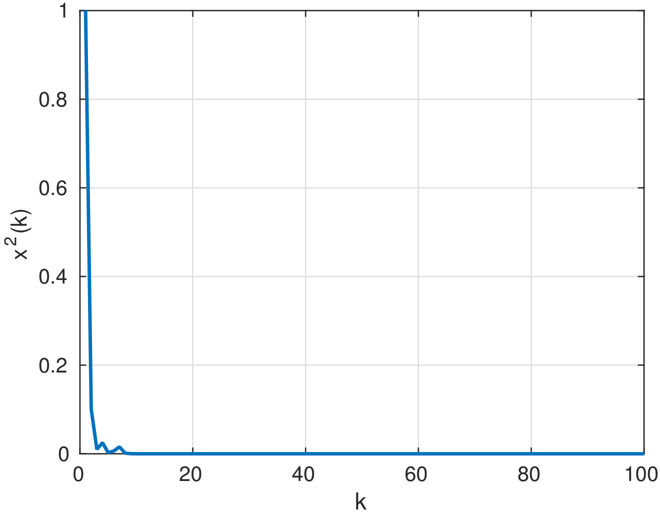
Understanding ASC

There is **zero probability** that we find a realisation of $\{x(k)\}_k$ which does not converge to 0.

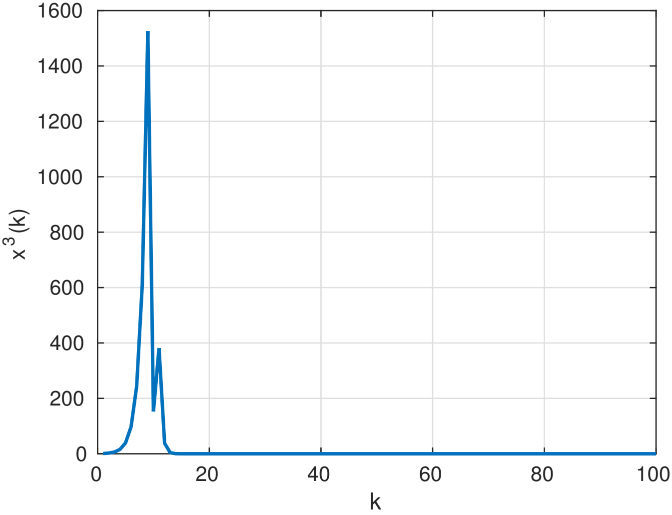
Understanding ASC



Understanding ASC



Understanding ASC



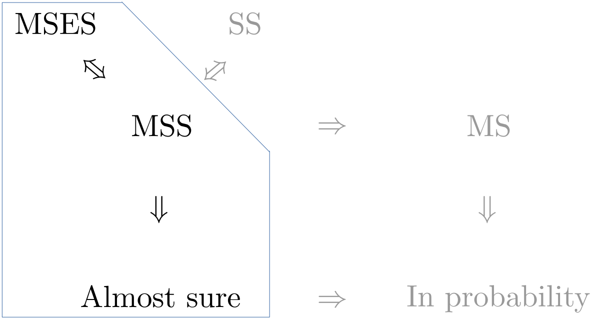
Modes of convergences

$$\begin{tikzcd}
\text{MSES} \arrow[dr, Leftrightarrow] & & \text{SS} \arrow[dl, Leftrightarrow] & \\
& \text{MSS} \arrow[rr, Rightarrow] \arrow[d, Rightarrow] & & \text{MS} \arrow[d, Rightarrow] \\
& \text{Almost sure} \arrow[rr, Rightarrow] & & \text{In probability}
\end{tikzcd}$$

Modes of convergences

$$\begin{tikzcd}
\text{MSES} & & \text{SS} & \\
& \text{MSS} \arrow[ul, Leftrightarrow] \arrow[ur, Leftrightarrow] \arrow[rr, Rightarrow] \arrow[d, Rightarrow] & & \text{MS} \arrow[d, Rightarrow] \\
& \text{Almost sure} \arrow[rr, Rightarrow] & & \text{In probability}
\end{tikzcd}$$

Modes of convergences



$$\begin{tikzcd}
\text{MSES} \arrow[dr, Leftrightarrow] & & \text{SS} \arrow[dl, Leftrightarrow] & \\
& \text{MSS} \arrow[rr, Rightarrow] \arrow[d, Rightarrow] & & \text{MS} \arrow[d, Rightarrow] \\
& \text{Almost sure} \arrow[rr, Rightarrow] & & \text{In probability}
\end{tikzcd}$$

Modes of convergences

$$\begin{tikzcd}
\text{MSES} & & \text{SS} & & \\
& \text{MSS} \arrow[ul, Leftrightarrow] \arrow[ur, Leftrightarrow] \arrow[rr, Rightarrow] \arrow[d, Rightarrow] & & \text{MS} \arrow[d, Rightarrow] \\
& \text{Almost sure} \arrow[rr, Rightarrow] & & \text{In probability}
\end{tikzcd}$$

Modes of convergences

$$\begin{tikzcd}
\text{MSES} \arrow[dr, Leftrightarrow] & & \text{SS} \arrow[dl, Leftrightarrow] & \\
& \text{MSS} \arrow[rr, Rightarrow] \arrow[d, Rightarrow] & & \text{MS} \arrow[d, Rightarrow] \\
& \text{Almost sure} \arrow[rr, Rightarrow] & & \text{In probability}
\end{tikzcd}$$

Question

Under what conditions $\|x(k)\| \to 0$ almost surely?

Interlude: Ergodicity

A finite Markov jump process $\{\theta(k)\}$ is called **ergodic** if from every mode $i \in \mathcal{N}$ we can move to any mode $j \in \mathcal{N}$ (not necessarily in one step).

Let $\pi_i(k) := \mathrm{P}\left[\theta(k) = i\right]$. Under the ergodicity assumption, there is a probability distribution $\pi = \{\pi_1, \ldots, \pi_N\}$, $\pi_i > 0$, such that

$$\pi_i(k) \to \pi_i.$$

And we can compute π solving[12]

$$P'\pi = \pi.$$

---

[12] For details, I recommend the lecture notes of J. Gravner (UC Davis) on "introductory probability".

Take ν pos. integer and let

$$\tilde{\theta}(k) = (\theta(k\nu + \nu - 1), \ldots, \theta(k\nu)) \in \mathcal{N}^{\nu}$$

be a Markov chain. Take $\iota = (i_{\nu-1}, \ldots, \iota_0)$ and $\tau = (j_{\nu-1}, \ldots, j_0) \in \mathcal{N}^{\nu}$. The transition probabilities of $\tilde{\theta}(k)$ are

$$\mathrm{P}\left[\tilde{\theta}(k+1) = \tau \mid \tilde{\theta}(k) = \iota\right] = p_{i_{\nu-1}j_0} p_{j_0 j_1} \ldots p_{j_{\nu-1}j_{\nu-1}}.$$

We have $\mathrm{P}\left[\tilde{\theta}(k) = \tau\right] = \mathrm{P}\left[\theta(k\nu) = j_0\right] p_{j_0 j_1} \ldots p_{j_{\nu-1}j_{\nu-1}}$ and as $k \to \infty$ this converges to

$$\tilde{\pi} := \pi_{j_0} p_{j_0 j_1} \ldots p_{j_{\nu-1}j_{\nu-1}}.$$

Conditions for ASC

Theorem

Let $\tilde{\Gamma}_\iota := \Gamma_{i_{\nu-1}} \dots \Gamma_{i_0}$. *If there is* ν *pos. integer. s.t.*

$$\prod_{\iota \in \mathcal{N}^\nu} \|\tilde{\Gamma}_\iota\|^{\tilde{\pi}_\iota} < 1,$$

then $x(k) \to 0$ *a.s.;* $\| \cdot \|$ *can be any matrix norm.*

Exercises

Exercise 1. Determine a system with $n = 2$, $N = 2$ which is ASC but not MSS.

Exercise 2. Show that a system with $n = 1$, $N = 2$, and $\theta(k)$: i.i.d. with $\pi_1 = \pi_2 = 0.5$ and $\Gamma_1 = \alpha$, $\Gamma_2 = \beta$ is ASC if and only if $|\alpha\beta| < 1$.

Exercise 3. For a stochastic process $\{x(k)\}_k$ we know that $\mathbb{E}[x(k)] \to 0$. Is the process ASC?

End of fifth section

Summary:

1. Almost sure convergence is weaker than MSS
2. ASC is convergence w.p.1.
3. ASC implies convergence in probability
4. The notion of almost sure convergence is not induced by a topology
5. A system can be ASC but not MSS or even MS
6. When the Markov process is ergodic, $\pi_i(k)$ converges

VI. Feedback stabilisation

Feedback stabilisation

We need to design a mean square stabilising controller for the system

$$x(k+1) = A_{\theta(k)}x(k) + B_{\theta(k)}u(k).$$

We assume $x(k)$, $\theta(k)$ are **measured** and the controller has the form

$$u(k) = \kappa(\theta(k), x(k)) = F_{\theta(k)}x(k),$$

so the closed-loop MJLS is

$$x(k+1) = (A_{\theta(k)} + B_{\theta(k)}F_{\theta(k)})x(k).$$

Feedback stabilisation

The closed-loop system is MSS if there is $V \in \mathbb{H}_{+}^{n}$ so that $\mathcal{T}(V) < V$, i.e., for all $j \in \mathcal{N}$

$$V_j - \sum_{i=1}^{N} p_{ij}(A_{\theta(k)} + B_{\theta(k)}F_{\theta(k)})V_i(A_{\theta(k)} + B_{\theta(k)}F_{\theta(k)})' > 0$$

which can be written as a **LMI** using the **Schur complement** lemma (Exercise).

Feedback stabilisation

Comments:

1. We have one controller for each mode.
2. For each $i \in \mathcal{N}$, we didn't select F_i so as to stabilise (A_i, B_i).
3. Had we done that, would the closed-loop MJLS be MSS?
4. In practice, $\theta(k)$ may not be known exactly.

Stabilisation with inexact knowledge of $\theta(k)$

Assume $\theta(k)$ is not available at time k, but we have an **estimate** $\hat{\theta}(k) \in \mathcal{N}$ (which defines a random process). Then,

$$u(k) = F_{\hat{\theta}(k)} x(k),$$

and the closed-loop MJLS will be

$$x(k+1) = (A_{\theta(k)} + B_{\theta(k)} F_{\hat{\theta}(k)}) x(k).$$

Stabilisation with inexact knowledge of $\theta(k)$

We define the following filtration $\{\mathfrak{M}_k\}_{k\in\mathbb{N}}$ with $\mathfrak{M}_0 = \mathcal{F}_0$ and $\mathfrak{M}_k$ is the σ-algebra generated by

$$\{x(0), \theta(0), x(1), \theta(1), \hat{\theta}(0), \ldots, x(k), \theta(k), \hat{\theta}(k-1)\}.$$

We assume that

$$\mathrm{P}\left[\hat{\theta}(k) = s \mid \mathfrak{M}_k\right] = \mathrm{P}\left[\hat{\theta}(k) = s \mid \theta(k)\right] =: \rho_{\theta(k)s}.$$

This means that

$$\rho_{is} = \{\text{probability that } \theta(k) = i \text{ while we estimated } \hat{\theta}(k) = s\}$$

and

$$\rho_{is} = \{\text{probability of success: } \theta(k) = \hat{\theta}(k) = i\}$$

Stabilisation with inexact knowledge of $\theta(k)$

Let

$$p = \min_i \rho_{ii}.$$

Let $\mathcal{T} : \mathbb{H}^n \to \mathbb{H}^n$ be defined as before with $\Gamma_i = A_i + B_i F_i$. Since $\mathcal{T}$ is stable, there are $\beta \geq 1$ and $\zeta \in (0, 1)$ s.t.[13]

$$\|\mathcal{T}^k\|_1 \leq \beta\zeta^k.$$

Finally, let

$$c_0 := \max\left\{\|A_i + B_i F_s\|^2, i, s \in \mathbb{N}, i \neq s, \rho_{is} > 0\right\}.$$

---

[13] Question: How do we determine those parameters? Can you give an example when $N = 1$ and $N = 2$?

Stabilisation with inexact knowledge of $\theta(k)$

MSS condition:

$$p > 1 - \frac{1-\zeta}{c_0 \beta}.$$

Proof.

The proof is a bit lengthy and can be found in the book of Costa *et al.* (Lemma 3.44). Sketch: We show that

$$0 \leq Q_j(k+1) \leq \mathcal{T}_j[Q(k)] + \mathcal{S}_j[Q(k)],$$

where $\mathcal{S} : \mathbb{H}^n \to \mathbb{H}^n$ is such that $\|\mathcal{S}\|_1 \leq c_0(1-p)$. ☐

Further reading

1. O.L.V. Costa, M.D. Fragoso and R.P. Marques, Discrete-time Markov Jump Linear Systems, Springer 2005.
2. Y. Fang and K.A. Loparo, Stochastic Stability of Jump Linear Systems, IEEE Trans. Aut. Contr. 47(7), pp. 1204–1208, 2002.

References for ASC

1. Y. Fang, K.A. Loparo, and X. Feng. Almost sure and δ-moment stability of jump linear systems. International Journal of Control, 59:1281–1307, 1994.
2. O. L. V. Costa and M. D. Fragoso. Stability results for discrete-time linear systems with Markovian jumping parameters. Journal of Mathematical Analysis and Applications, 179:154–178, 1993.
3. J. B. R. do Val, C. Nespoli, and Y. R. Z. Caceres. Stochastic stability for Markovian jump linear systems associated with a finite number of jump times. Jounal of Mathematical Analysis and Applications, 285:551–563, 2003.
4. Y. Fang. A new general sufficient condition for almost sure stability of jump linear systems. IEEE Transactions on Automatic Control, 42:378–382, 1997.
5. Z. G. Li, Y. C. Soh, and C. Y. Wen. Sufficient conditions for almost sure stability of jump linear systems. IEEE Transactions on Automatic Control, 45:1325–1329, 2000.
6. M. Mariton. Almost sure and moments stability of jump linear systems. Systems and Control Letters, 30:1145–1147, 1985.