Mean-field methods: what can go wrong?

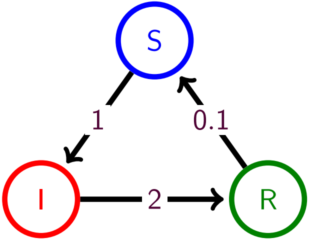
The decoupling assumption: a zoom on the fixed point and on mean-field games

Nicolas Gast (Inria) and Luca Bortolussi (UNITS)

Inria, Grenoble, France

SFM, Bertinoro, June 21, 2016

Markov chains

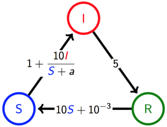


Transition graph

$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$

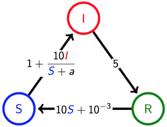
Infinitesimal generator

Transient and steady-state analysis



$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$

Transient and steady-state analysis



$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$

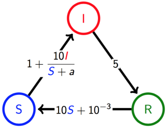
Transient analysis: the master equation

If X is a CTMC (continuous time Markov chain) with generator Q:

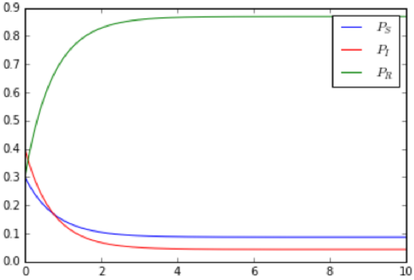
$$\frac{d}{dt}P_i(t) = \sum_{j \in S} P_j(t) Q_{ji},$$

where $P_i(t) = \mathbb{P}(X(t) = i)$.

Transient and steady-state analysis



$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$



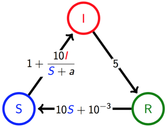
Transient analysis: the master equation

If X is a CTMC (continuous time Markov chain) with generator Q:

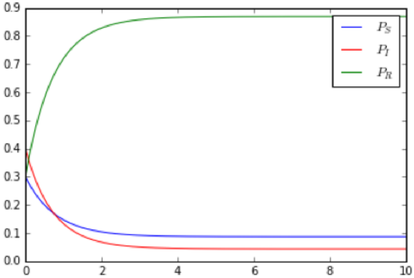
$$\frac{d}{dt}P(t) = P(t)Q_{ji},$$

where $P_i(t) = \mathbb{P}(X(t) = i)$.

Transient and steady-state analysis

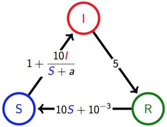


$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$

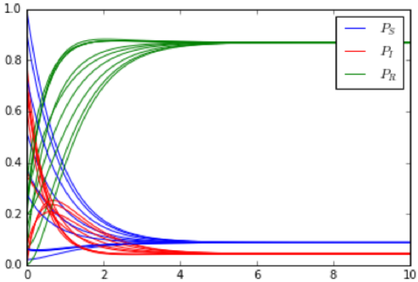


Steady-state analysis

Transient and steady-state analysis



$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0.1 & 0 & -0.1 \end{pmatrix}$$

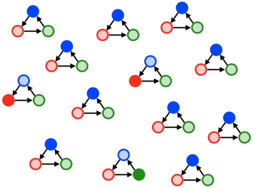


Steady-state analysis

If the chain is irreducible,

- The equation $\pi Q = 0$ has a unique solution such that $\sum_i \pi_i = 1$.
- $\lim_{i \to \infty} P_i(t) = \pi_i$

The state space explosion



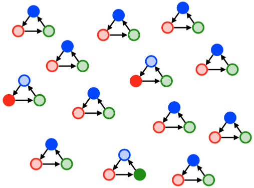
$3^{13} \approx 10^6$ states.

We need to keep track of S^N states

$$\mathbb{P}(X_1(t) = i_1, \ldots, X_n(t) = i_n)$$

The generator Q has S^N entries.

The state space explosion



$3^{13} \approx 10^6$ states.

We need to keep track of S^N states

$$\mathbb{P}(X_1(t) = i_1, \dots, X_n(t) = i_n)$$

The generator Q has S^N entries.

The decoupling assumption is

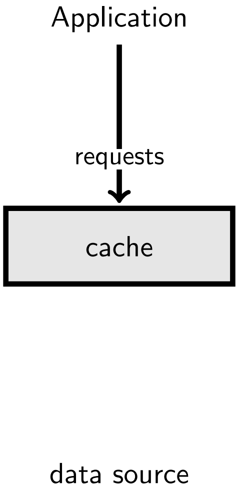
$$\underbrace{\mathbb{P}(X_1(t) = i_1, \dots, X_n(t) = i_n)}_{S^N \text{ variables}} \approx \underbrace{\mathbb{P}(X_1(t) = i_1) \dots \mathbb{P}(X_n(t) = i_n)}_{N \times S \text{ variables}}$$

Question: when is this (not) valid?

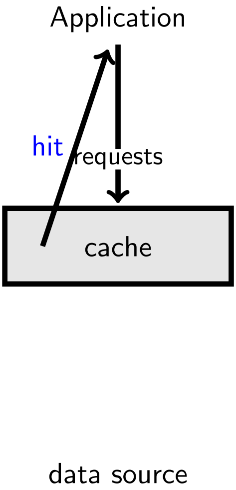
Outline

Outline

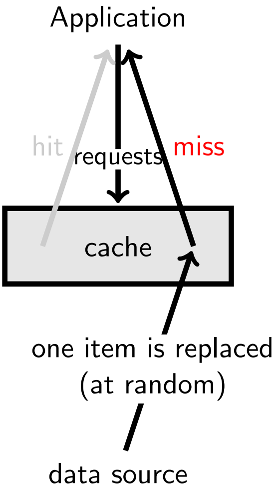
A cache-replacement policy
G. Van Houdt, 2015



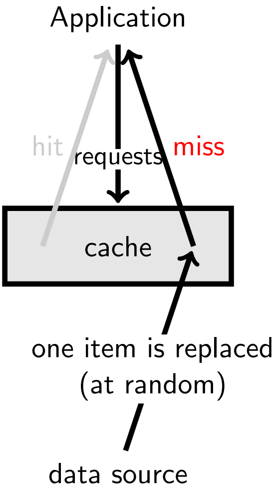
A cache-replacement policy
G. Van Houdt, 2015



A cache-replacement policy
G. Van Houdt, 2015



A cache-replacement policy
G. Van Houdt, 2015



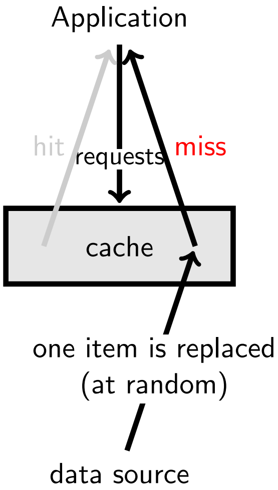
Model:

- Items have the same size.
- Cache can store m items.
- There are n items. Item i is requested with probability p_i.

Goal

- Compute $\mathbb{P}(\text{item 1 is in cache})$
- Compute hit probability.

A cache-replacement policy
G. Van Houdt, 2015



Markov model

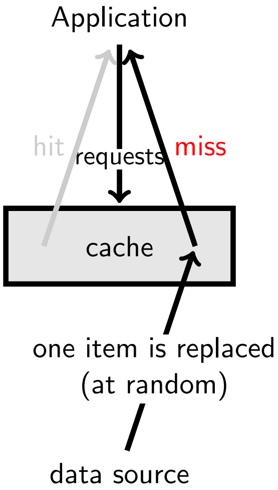
State space : set of m distinct items.

Transitions:

$$\{i_1 \dots i_m\} \mapsto \{i_1 \dots i_{k-1}, j, i_{k+1} \dots i_n\}$$

with probability p_j/m.

A cache-replacement policy
G. Van Houdt, 2015



Markov model

State space : set of m distinct items.

Transitions:

$$\{i_1 \dots i_m\} \mapsto \{i_1 \dots i_{k-1}, j, i_{k+1} \dots i_n\}$$

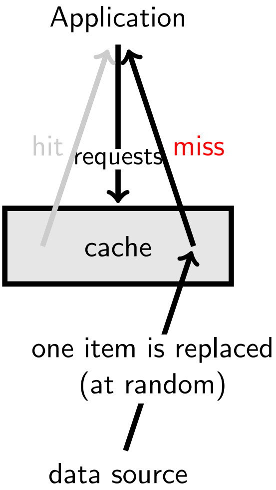
with probability p_j/m.

Decoupling assumption

$$\mathbb{P}(i_1 \dots i_m) \approx \underbrace{\mathbb{P}(i_1)}_{=:x_{i_1}} \dots \mathbb{P}(i_m)$$

A cache-replacement policy

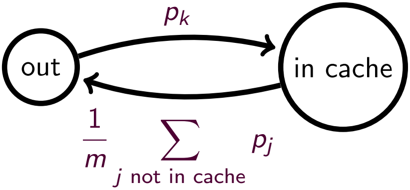
G. Van Houdt, 2015



Decoupling assumption

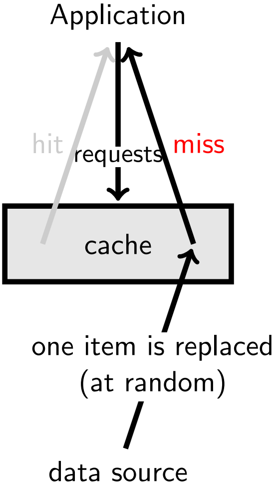
$$\mathbb{P}(i_1 \ldots i_m) \approx \underbrace{\mathbb{P}(i_1)}_{=:x_{i_1}} \ldots \mathbb{P}(i_m)$$

If we zoom on object k:



A cache-replacement policy

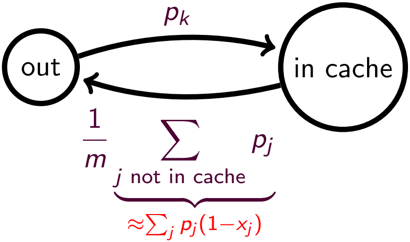
G. Van Houdt, 2015



Decoupling assumption

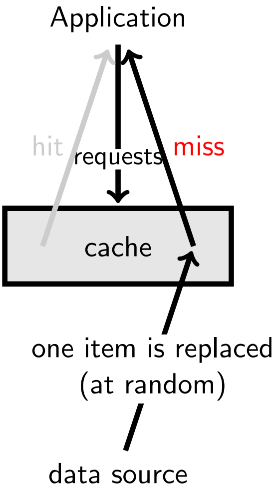
$$\mathbb{P}(i_1 \dots i_m) \approx \underbrace{\mathbb{P}(i_1)}_{=:x_{i_1}} \dots \mathbb{P}(i_m)$$

If we zoom on object k:

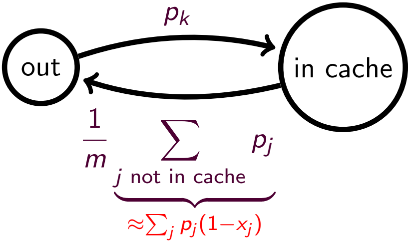


A cache-replacement policy

G. Van Houdt, 2015



If we zoom on object k:

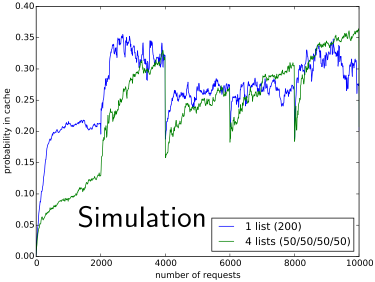


Mean-field model

Let $x_k := \mathbb{P}(\text{item } k \text{ is in the cache})$.

$$\dot{x}_k = p_k(1 - x_k) - \frac{\sum_\ell (p_\ell(1-x_\ell))}{m} x_k.$$

A cache-replacement policy: simulation



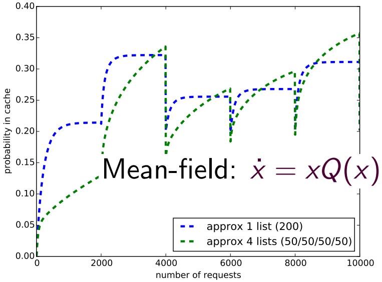


Figure: Popularities of objects change every 2000 steps.

A cache-replacement policy: simulation

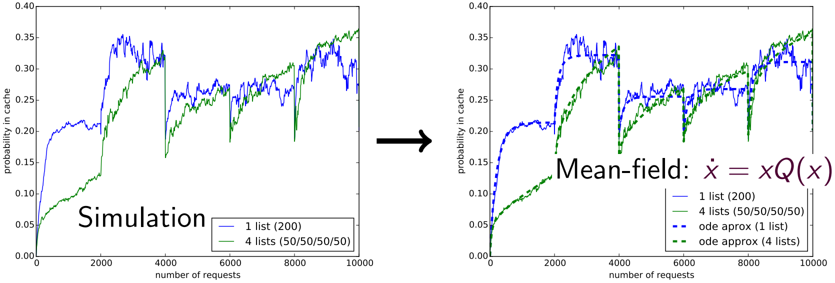


Figure: Popularities of objects change every 2000 steps.

Stationary distribution

Fixed point equation

- $0 = \dot{x}_k = p_k(1 - x_k) - \frac{\sum_\ell (p_\ell (1 - x_\ell))}{m} x_k.$
- $\sum_k x_k = m.$

(ref: Dan and Towsley, Gast Van Houdt, ...)

Stationary distribution

Fixed point equation

- $0 = \dot{x}_k = p_k(1 - x_k) - \frac{\sum_\ell (p_\ell(1-x_\ell))}{m} x_k$.
- $\sum_k x_k = m$.

(ref: Dan and Towsley, Gast Van Houdt, ...)

Algorithm: easy to solve:

1. Define $x_k(T)$ the solution of $p_k(1 - x_k) - Tx_k$.
   - $x_k(T) = p_k/(1 + T)$
2. Find T such that $\sum_k (1 - x_k(T)) = m$.

Outline

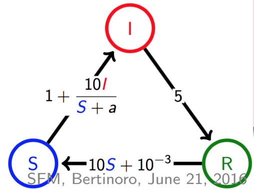
Decoupling and $\dot{x} = xQ(x)$

$$\mathbb{P}(X_1(t) = i_1, \dots, X_n(t) = i_n) \approx \underbrace{\mathbb{P}(X_1(t) = i_1)}_{=x_{1,i_1}(t)} \dots \underbrace{\mathbb{P}(X_n(t) = i_n)}_{=x_{n,i_n}(t)}$$

When we zoom on one object

$$\mathbb{P}(X_1(t+dt) = j | X_1(t) = i) \approx \mathbb{E}\left[\mathbb{P}(X_1(t) = j | X_1 = i \wedge X_2 \dots X_n)\right]$$
$$\approx Q^{(1)}_{i,j}(\mathbf{x}) := \sum_{i_2 \dots i_n} K_{(i,i_2\dots i_n)\to(j,j_2\dots j_n)} x_{2,i_2} \dots x_{n,i_n}$$

We then get: $\frac{d}{dt} x_{1,j}(t) \approx \sum_i x_{1,i} Q^{(1)}_{i,j}$



Transient regime

Theorem (Snitzman (99), Kurtz (70'), Benaim, Le Boudec (08),...)

For fixed t, the decoupling assumption is equivalent to the mean-field convergence.

For example (remember Luca's talk), if $x \mapsto xQ(x)$ is Lipschitz-continuous then, as the number of objects N goes to infinity:

$$\lim_{N\to\infty} \mathbb{P}(X_k(t) = i) = x_{k,i}(t),$$

where x satisfies $\dot{x} = xQ(x)$.

The fixed point method

Markov chain

Transient regime $\dot{p} = pK$

$$\downarrow t \to \infty$$

Stationary $\pi K = 0$

[1] Performance analysis of the IEEE 802.11 distributed coordination function.

[2] Fixed point analys is of single cell IEEE 802.11e WLANs: Uniqueness, multistability.

[3] Performance analysis of exponenetial backoff.

[4] New insights from a fixed-point analysis of single cell IEEE 802.11 WLANs.

The fixed point method

$$
\begin{tikzcd}[column sep=large]
 & \text{Markov chain} & & \text{Mean-field} \\
\text{Transient regime} & \dot{p} = pK \arrow[rr, "N \to \infty"] \arrow[d, "t \to \infty"] & & \dot{x} = xQ(x) \\
\text{Stationary} & \pi K = 0 \arrow[rr, "?"] & & \boxed{\substack{xQ(x) = 0 \\ \text{fixed points}}}
\end{tikzcd}
$$

Method was used in many papers: Bianchi 00[1] Ramaiyan et al. 08[2] Kwak et al. 05[3] Kumar et al 08[4]

[1] Performance analysis of the IEEE 802.11 distributed coordination function.

[2] Fixed point analys is of single cell IEEE 802.11e WLANs: Uniqueness, multistability.

[3] Performance analysis of exponenetial backoff.

[4] New insights from a fixed-point analysis of single cell IEEE 802.11 WLANs.

Does it always work?[5][6]

SIRS model:

- A node S becomes I at rate 1 (external infection)
- When a S meets an I, it becomes infected at rate $1/(S+a)$
- An I recovers at rate 5.
- A node R becomes S by:
  - meeting a node S (rate $10S$)
  - alone (at rate 10^{-3}).

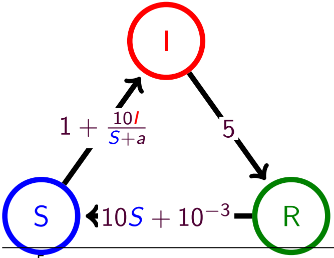
[5] Benaim Le Boudec 08

[6] Cho, Le Boudec, Jiang, On the Asymptotic Validity of the Decoupling Assumption for Analyzing 802.11 MAC Protoco. 2010

Does it always work?[5][6]

SIRS model:

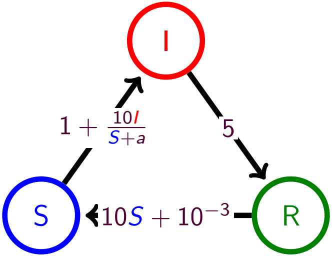
- A node S becomes I at rate 1 (external infection)
- When a S meets an I, it becomes infected at rate $1/(S+a)$
- An I recovers at rate 5.
- A node R becomes S by:
  - meeting a node S (rate $10S$)
  - alone (at rate 10^{-3}).



[5] Benaim Le Boudec 08

[6] Cho, Le Boudec, Jiang, On the Asymptotic Validity of the Decoupling Assumption for Analyzing 802.11 MAC Protoco. 2010

Does it always work?[7,8]

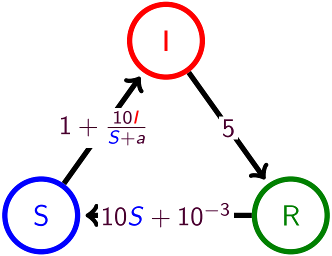


- Markov chain is irreducible.
- Unique fixed point $xQ(x) = 0$.

[7] Benaim Le Boudec 08

[8] Cho, Le Boudec, Jiang, On the Asymptotic Validity of the Decoupling Assumption for Analyzing 802.11 MAC Protoco. 2010

Does it always work?[7][8]



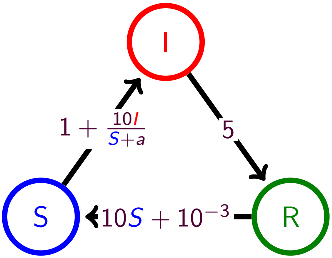
- Markov chain is irreducible.
- Unique fixed point $xQ(x) = 0$.

	Fixed point $xQ(x) = 0$		Stat. measure $N = 1000$	
	x_S	x_I	π_S	π_I
$a = .3$	0.209	0.234	0.209	0.234

[7] Benaim Le Boudec 08

[8] Cho, Le Boudec, Jiang, On the Asymptotic Validity of the Decoupling Assumption for Analyzing 802.11 MAC Protoco. 2010

Does it always work?[7][8]



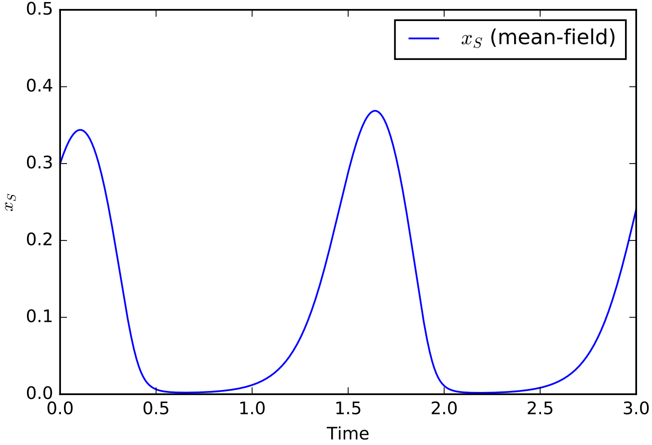
- Markov chain is irreducible.
- Unique fixed point $xQ(x) = 0$.

	Fixed point $xQ(x) = 0$		Stat. measure $N = 1000$	
	x_S	x_I	π_S	π_I
$a = .3$	0.209	0.234	0.209	0.234
$a = .1$	0.078	0.126	0.11	0.13

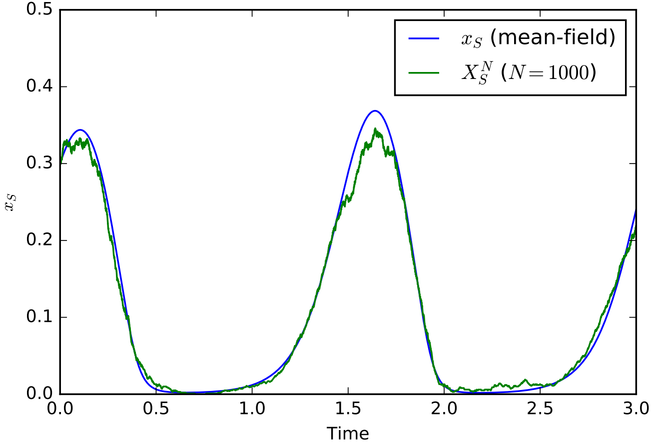
[7] Benaim Le Boudec 08

[8] Cho, Le Boudec, Jiang, On the Asymptotic Validity of the Decoupling Assumption for Analyzing 802.11 MAC Protoco. 2010

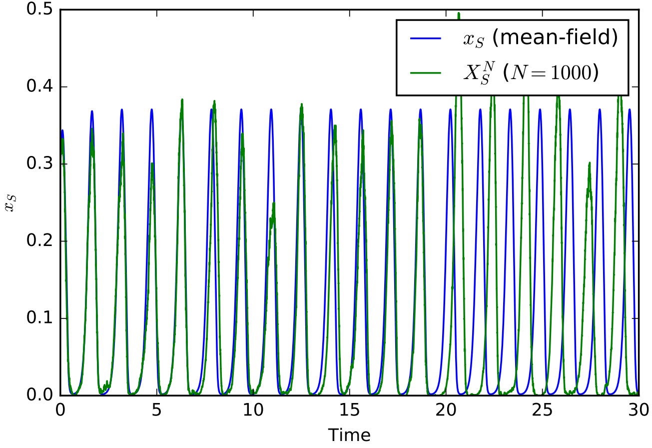
What happened?



What happened?

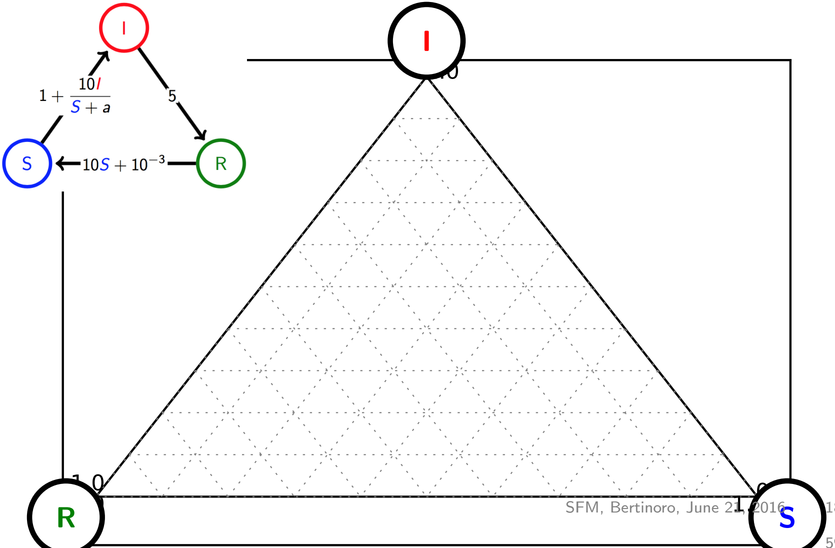


What happened?



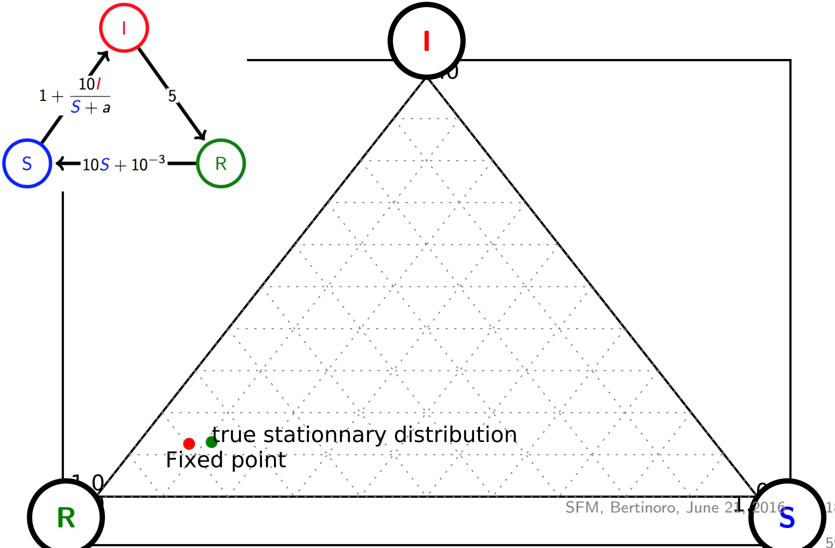
What happened?

$(x_S = 0.078, x_I = 0.126), (\pi_S = 0.11, \pi_I = 0.13)$



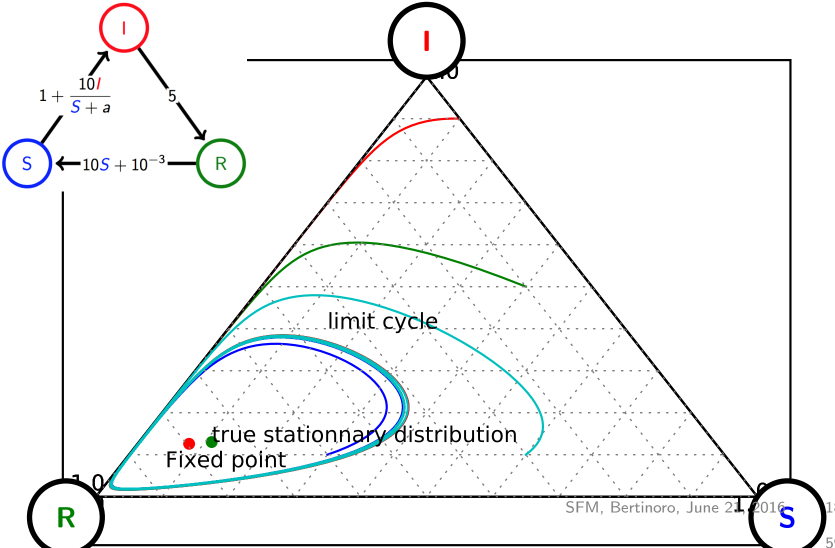
What happened?

$(x_S = 0.078, x_I = 0.126), (\pi_S = 0.11, \pi_I = 0.13)$



What happened?

$(x_S = 0.078, x_I = 0.126), (\pi_S = 0.11, \pi_I = 0.13)$



Fixed points?

Markov chain

Transient regime $\dot{p} = pK$

$$\Big\downarrow t \to \infty$$

Stationary $\pi K = 0$

Fixed points?

	Markov chain		Mean-field
Transient regime	$\dot{p} = pK$	$\xrightarrow{N \to \infty}$	$\dot{x} = xQ(x)$
	$\big\downarrow t \to \infty$		
Stationary	$\pi K = 0$	$\xrightarrow{?}$	$\boxed{xQ(x) = 0}$ fixed points

Fixed points?

	Markov chain		Mean-field
Transient regime	$\dot{p} = pK$	$\xrightarrow{N \to \infty}$	$\dot{x} = xQ(x)$
	$\downarrow t \to \infty$		$\downarrow \cancel{t \to \infty}$
Stationary	$\pi K = 0$	$\xrightarrow{\cancel{N \to \infty}}$	$\boxed{\cancel{xQ(x) = 0}}$ fixed points

Fixed points?

	Markov chain		Mean-field
Transient regime			
Stationary			

$$\begin{tikzcd}
\dot{p} = pK \arrow[rr, "N \to \infty"] \arrow[d, "t \to \infty"] & & \dot{x} = xQ(x) \arrow[d, "\cancel{t \to \infty}"] \arrow[d, bend left, "\text{if yes}"] \\
\pi K = 0 \arrow[rr, "\cancel{N \to \infty}"] \arrow[rr, bend right, "\text{then yes}"'] & & \boxed{xQ(x) = 0 \text{ fixed points}}
\end{tikzcd}$$

Theorem ((i) Benaim Le Boudec 08,(ii) Le Boudec 12)

The stationary distribution π^N concentrates on the fixed points if :

(i) *All trajectories of the ODE converges to the fixed points.*

(ii) *(or) The Markov chain is reversible.*

Steady-state: theorem

Theorem

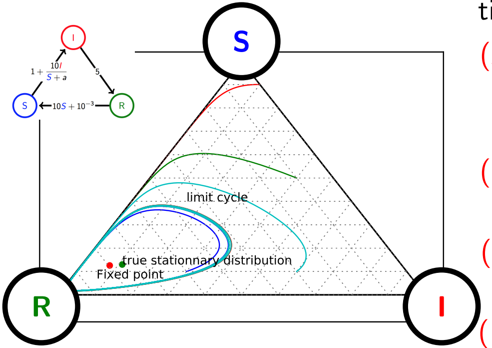
Let us consider a mean-field model for which x^N converges to the solution of $\dot{x} = f(x)$. Then:

- *If all trajectories converge to a unique fixed point x^*, the π^N converges to x^*.*

Note: unique fixed point implies the decoupling assumption:

Quiz

Consider the SIRS model:



Under the stationary distribution π^N:

(A) As there are no fixed point, there is no such stationary distribution.

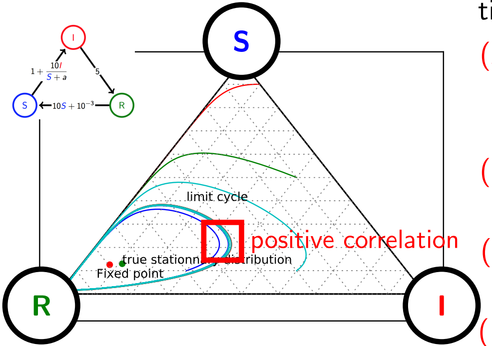
(B) $P(X_1 = S, X_2 = S) \approx P(X_1 = S)P(X_2 = S)$

(C) $P(X_1 = S, X_2 = S) > P(X_1 = S)P(X_2 = S)$

(D) $P(X_1 = S, X_2 = S) < P(X_1 = S)P(X_2 = S)$

Quiz

Consider the SIRS model:



Under the stationary distribution π^N:

(A) As there are no fixed point, there is no such stationary distribution.

(B) $P(X_1 = S, X_2 = S) \approx P(X_1 = S)P(X_2 = S)$

(C) $P(X_1 = S, X_2 = S) > P(X_1 = S)P(X_2 = S)$

(D) $P(X_1 = S, X_2 = S) < P(X_1 = S)P(X_2 = S)$

Answer: C

$P(X_1(t) = S, X_2(t) = S) = x_1(t)^2$. Thus: positively correlated.

Lyapunov functions

How to show that trajectories converge to a fixed point?

Lyapunov functions

How to show that trajectories converge to a fixed point?

A solution of $\frac{d}{dt}x(t) = xQ(x(t))$ converges to the fixed points of $xQ(x) = 0$, if there exists a Lyapunov function f, that is:

- Lower bounded: $\inf_x f(x) > +\infty$
- Decreasing along trajectories:

$$\frac{d}{dt}f(x(t)) < 0,$$

whenever $x(t)Q(x(t)) \neq 0$.

A solution of $\frac{d}{dt}x(t) = xQ(x(t))$ converges to the fixed points of $xQ(x) = 0$, if there exists a Lyapunov function f, that is:

- Lower bounded: $\inf_x f(x) > +\infty$
- Decreasing along trajectories:

$$\frac{d}{dt}f(x(t)) < 0,$$

whenever $x(t)Q(x(t)) \neq 0$.

How to find a Lyapunov function

- Energy? Distance? Entropy? Luck?

The relative entropy is a Lyapunov function for Markov chains

Let Q be the generator of an irreducible Markov chain and π be its stationary distribution. Let $P(t)$ be the solution of $\frac{d}{dt}P(t) = P(t)Q$.

Theorem (e.g. Budhiraja et al 15, Dupuis-Fischer 11)

The relative entropy

$$R(P\|\pi) = \sum_i P_i \log \frac{P_i}{\pi_i}$$

is a Lyapunov function:

$$\frac{d}{dt}R(P(t)\|\pi) < 0,$$

with equality if and only if $P(t) = \pi$.

Relative entropy for mean-field models

Assume that $Q(x)$ be a generator of an irreducible Markov chain and let $\pi(x)$ be its stationary distribution. Let $P(t)$ be the solution of $\frac{d}{dt}P(t) = P(t)Q(P(t))$. Then

$$\frac{d}{dt}R(P(t)\|\pi(t)) = \underbrace{\frac{d}{dt}P(t)\frac{\partial}{\partial P}R(P(t),\pi(t))}_{\leq 0} + \underbrace{\frac{d}{dt}\pi(t)\frac{\partial}{\partial \pi}R(P(t),\pi(t))}_{=-\sum_i x_i(t)\frac{d}{dt}\log\pi_i(t)}$$

$$\leq -\sum_i x_i(t)\frac{d}{dt}\log\pi_i(t)$$

Relative entropy for mean-field models

Assume that $Q(x)$ be a generator of an irreducible Markov chain and let $\pi(x)$ be its stationary distribution. Let $P(t)$ be the solution of $\frac{d}{dt}P(t) = P(t)Q(P(t))$. Then

$$\frac{d}{dt}R(P(t)\|\pi(t)) = \underbrace{\frac{d}{dt}P(t)\frac{\partial}{\partial P}R(P(t),\pi(t))}_{\leq 0} + \underbrace{\frac{d}{dt}\pi(t)\frac{\partial}{\partial \pi}R(P(t),\pi(t))}_{=-\sum_i x_i(t)\frac{d}{dt}\log \pi_i(t)}$$

$$\leq -\sum_i x_i(t)\frac{d}{dt}\log \pi_i(t)$$

Theorem

If there exists a lower bounded integral $F(x)$ of $-\sum_i x_i(t)\frac{d}{dt}\log \pi_i(t)$, then $x \mapsto R(x\|\pi(x)) + F(x)$ is a Lyapunov function for the mean-field model.

The decoupling assumption: conclusion

- Decoupling $\approx$ mean-field convergence
- If the rates are continuous, convergence holds for the transient regime
- The stationary regime should be handle with care
  - The uniqueness of the fixed point is not enough.
  - Lyapunov functions can help but are not easy to find.

Outline

A martingale argument

The drift of a mean-field model is $X(t)$ satisfies

$$\lim_{dt\to 0} \frac{1}{dt}\mathbb{E}\left[X(t+dt) - X(t) | X(t) = x\right] = f(x)$$

$$\lim_{dt\to 0} \frac{1}{dt}\mathrm{var}\left[X(t+dt) - X(t) - f(X(t)) | X(t) = x\right] \leq C/N$$

This means that:

$$M(t) = X(t) - (x_0 - \int_0^t f(X(s))ds)$$

is such that:

$$\underbrace{\mathbb{E}\left[M(t) \mid \mathcal{F}_s\right] = M(s)}_{M(t) \text{ is a martingale}} \qquad \wedge \qquad \underbrace{\mathrm{var}\left[M(t)\right] \leq Ct/N}_{\text{Small variance}}.$$

Martingale concentration results

Let $M(t)$ be such that:

$$\underbrace{\mathbb{E}\left[M(t) \mid \mathcal{F}_s\right] = M(s)}_{M(t) \text{ is a martingale}} \qquad \wedge \qquad \underbrace{\mathrm{var}\left[M(t)\right] \leq C/N}_{\text{Small variance}}.$$

Then: (Doob's inequality):

$$\mathbf{P}\left[\sup_{t\leq T} \|M(t)\| \geq \epsilon\right] \leq \frac{C}{N\epsilon^2}.$$

Mean-field convergence

Going back to slide 1, we have:

$$X(t) = x_0 + \int_0^t f(X(s))ds + \underbrace{M(t)}_{\text{small by previous slide}}$$

Mean-field convergence

Going back to slide 1, we have:

$$X(t) = x_0 + \int_0^t f(X(s))ds + \underbrace{M(t)}_{\text{small by previous slide}}$$

Is $X(t)$ close to $\dot{x} = f(x)$?

The initial value problem
"Dynamical systems 101"

The initial value problem:

$$\begin{cases} \dot{x} = f(x) \\ x(0) = x_0 \in \mathbb{R}^d. \end{cases}$$

The existence and solution is guaranteed by the Picard-Cauchy theorem:

- If f is Lipschitz-continuous on $\mathbb{R}^d$, then there exists a unique solution on $[0, T]$.

Uniqueness of solution

"Dynamical system 101 (ctn)"

Reminder: f is Lipschitz-continuous if there exists L such that:
$\forall x, y \in \mathbb{R}^d$:

$$\|f(x) - f(y)\| \leq L \|x - y\| .$$

Uniqueness of solution

"Dynamical system 101 (ctn)"

Reminder: f is Lipschitz-continuous if there exists L such that: $\forall x, y \in \mathbb{R}^d$:

$$\|f(x) - f(y)\| \leq L \|x - y\| .$$

If $x(t) = x_0 + \int_0^t f(x(s))ds$ and $y(t) = y_0 + \int_0^t f(y(s))ds + \varepsilon$ then

$$\|x(t) - y(t)\| \leq L \int_0^t \|x(s) - y(s)\| + \|x_0 - y_0\| + \varepsilon.$$

Uniqueness of solution
"Dynamical system 101 (ctn)"

Reminder: f is Lipschitz-continuous if there exists L such that: $\forall x, y \in \mathbb{R}^d$:

$$\|f(x) - f(y)\| \leq L \|x - y\|.$$

If $x(t) = x_0 + \int_0^t f(x(s))ds$ and $y(t) = y_0 + \int_0^t f(y(s))ds + \varepsilon$ then

$$\|x(t) - y(t)\| \leq L \int_0^t \|x(s) - y(s)\| + \|x_0 - y_0\| + \varepsilon.$$

Gronwall's Lemma: this implies that

$$\|x(t) - y(t)\| \leq (\|x_0 - y_0\| + \varepsilon)e^{Lt}.$$

Consequence

Theorem

If $X^N(0) = x_0$, then:

$$\mathbb{E}\left[\sup_{t \leq T} \left\| X^N(t) - x(t) \right\|\right] \leq O\left(\frac{1}{\sqrt{N}}\right) e^{LT}.$$

Rate of convergence: recap and some extensions

The speed of convergence can be extended to

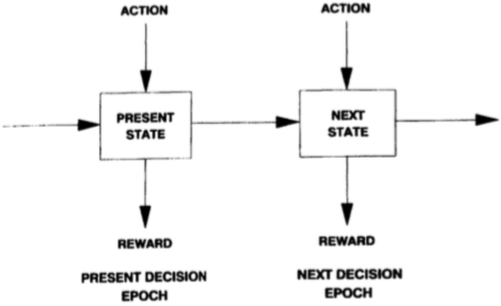
- Non-smooth dynamics (one sided Lipschitz functions)
- Steady-state (if f is C^2 and unique attractor)
- $\mathbb{E}[X(t)]$

It cannot be extended to

- General non-Lipschitz dynamics.
- Steady-state with no attractor.

Outline

Optimal control



- Stochastic optimal control: closed-loop policies actions(t+1)=function(state(t)).
- Deterministic optimal control: open-loop policies are optimal.

Markov decision processes

Reference: Puterman (2014)

Definition: a Markov decision process (MDP)

- State space / action space
- Transition probabilities : $p(X(t+1) = j | X(t) = i, action)$
- Instantaneous cost: cost(t, state, action).
- Objective:

$$\min \mathbb{E}\left[\mathrm{cost}(t, X_t, action)\right]$$

Example: You can throw a 6-face dice up to 5 times. You win the number on the last dice. When should you stop?

Definition: a Markov decision process (MDP)

- State space $\{1 \dots 6\}$ / action space $=\{$stop, continue$\}$
- Transition probabilities : $p(X(t+1)=j|X(t)=i, action)$
  $p(X(t+1)=i)=1/6$ if continue. $p(X(t+1)=X(t))=1$ if stop.
- Instantaneous cost: cost(t, state, action).
- Objective:

$$\min \mathbb{E}\left[\mathrm{cost}(t, X_t, action)\right]$$

Example of Markov decision process

You can throw a 6-face dice up to 5 times. You win the number on the last dice. When should you stop?
Value iteration (Bellman's equation)

$$V_t(i) = \max_{action} \mathrm{cost}(t, i, action) + \mathbb{E}\left[V_{t+1}(X(t+1) \mid X(t) = i, action)\right].$$

Example:

t / i	1	2	3	4	5
1					
2					
3					
4					
5					
6					

Example of Markov decision process

You can throw a 6-face dice up to 5 times. You win the number on the last dice. When should you stop?
Value iteration (Bellman's equation)

$$V_t(i) = \max_{action} \mathrm{cost}(t, i, action) + \mathbb{E}\left[V_{t+1}(X(t+1) \mid X(t) = i, action)\right].$$

Example:

i \ t	1	2	3	4	5
1					1
2					2
3					3
4					4
5					5
6					6

Example of Markov decision process

You can throw a 6-face dice up to 5 times. You win the number on the last dice. When should you stop?
Value iteration (Bellman's equation)

$$V_t(i) = \max_{action} \mathrm{cost}(t, i, action) + \mathbb{E}\left[V_{t+1}(X(t+1) \mid X(t) = i, action)\right].$$

Example:

t / i	1	2	3	4	5
1				3.5	1
2				3.5	2
3				3.5	3
4				4	4
5				5	5
6				6	6

Example of Markov decision process

You can throw a 6-face dice up to 5 times. You win the number on the last dice. When should you stop?
Value iteration (Bellman's equation)

$$V_t(i) = \max_{action} \mathrm{cost}(t, i, action) + \mathbb{E}\left[V_{t+1}(X(t+1) \mid X(t) = i, action)\right].$$

Example:

i \ t	1	2	3	4	5
1			4.25	3.5	1
2			4.25	3.5	2
3			4.25	3.5	3
4			4.25	4	4
5			5	5	5
6			6	6	6

Example of Markov decision process

You can throw a 6-face dice up to 5 times. You win the number on the last dice. When should you stop?
Value iteration (Bellman's equation)

$$V_t(i) = \max_{action} \mathrm{cost}(t, i, action) + \mathbb{E}\left[V_{t+1}(X(t+1) \mid X(t) = i, action)\right].$$

Example:

i \ t	1	2	3	4	5
1		4.66	4.25	3.5	1
2		4.66	4.25	3.5	2
3		4.66	4.25	3.5	3
4		4.66	4.25	4	4
5		5	5	5	5
6		6	6	6	6

Example of Markov decision process

You can throw a 6-face dice up to 5 times. You win the number on the last dice. When should you stop?
Value iteration (Bellman's equation)

$$V_t(i) = \max_{action} \mathrm{cost}(t, i, action) + \mathbb{E}\left[V_{t+1}(X(t+1) \mid X(t) = i, action)\right].$$

Example:

i \ t	1	2	3	4	5
1	4.95	4.66	4.25	3.5	1
2	4.95	4.66	4.25	3.5	2
3	4.95	4.66	4.25	3.5	3
4	4.95	4.66	4.25	4	4
5	5	5	5	5	5
6	6	6	6	6	6

The curse of dimensionality

To solve Bellman's equation, we need to iterate over the whole state space.

$$V_t(i) = \min_{action} \mathrm{cost}(t, i, action) + \mathbb{E}\left[V_{t+1}(X(t+1) \mid X(t) = i, action)\right].$$

The curse of dimensionality

To solve Bellman's equation, we need to iterate over the whole state space.

$$V_t(i) = \min_{action} \mathrm{cost}(t, i, action) + \mathbb{E}\left[V_{t+1}(X(t+1) \mid X(t) = i, action)\right].$$

Alternative:

- Approximate dynamic programming (learning)
- Mean-field optimal control

Example of mean-field control

MDP	Mean-field optimization
Find $\pi(t, X)$ to minimize $$V^{\pi,N} = \mathbb{E}\left[\sum_t cost(X_t, \pi(t, X_t))\right]$$ subject to $P(X_{t+1} = i \mid X_t = j, \pi(.) = a) = P_{i,j,a}$.	Find $a(t)$ to minimize $$V^a = \int_0^T cost(x_t, a_t)dt$$ subject to $\dot{x}_t = f(x_t, a_t)$

Example of mean-field control

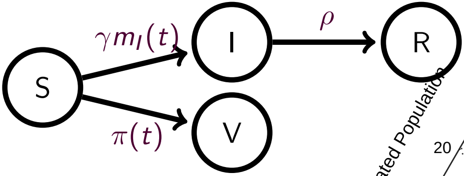
MDP	Mean-field optimization
Find $\pi(t, X)$ to minimize $$V^{\pi,N} = \mathbb{E}\left[\sum_t cost(X_t, \pi(t, X_t))\right]$$ subject to $P(X_{t+1} = i \mid X_t = j, \pi(.) = a) = P_{i,j,a}$.	Find $a(t)$ to minimize $$V^a = \int_0^T cost(x_t, a_t)dt$$ subject to $\dot{x}_t = f(x_t, a_t)$

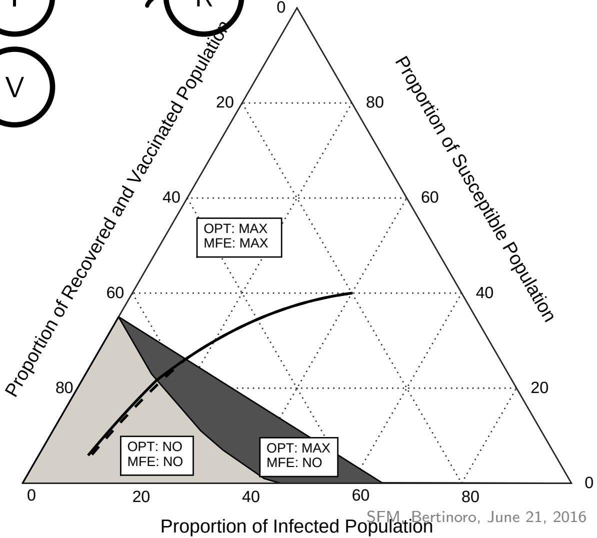
Theorem (G. Gaujal, Le Boudec 2012)

If the drift and costs are Lipschitz, then

- *the $V^{N,*} \to V^*$*
- *An open-loop policy a^* is optimal*

Mean-field control: example





Outline

Motivation

Mean field games (Lions and Lasry, 2007 and Caines, 2007) capture the dynamic evolution of a large population of strategic players.

Game Taxinomy

- static games: payoff matrix per player.
  Strategy of one player is a (randomized) action.
- Stochastic (repeated) games: payoff is the (disc.) sum from 0 to T.
  Strategy of a player is a policy (function).
- population games: infinite number of identical players.
  Players profiles replaced by action profiles.
- Mean field games: dynamic games over infinite number of players.

Game Taxinomy

- static games: payoff matrix per player.
  Strategy of one player is a (randomized) action.
  Solution of the game: Nash equilibrium.

- population games: infinite number of identical players.
  Players profiles replaced by action profiles.
  Solution of the game: Wardrop equilibrium

- Stochastic (repeated) games: payoff is the (disc.) sum from 0 to T.
  Strategy of a player is a policy (function).
  Solution: Sub-game Perfect Eq. + folk theorem.

- Mean field games: dynamic games over infinite number of players.
  Solution of the game: mean field equilibrium.

Static game example
The prisoner's dilemma

Two possible actions: $\{C, D\}$.
The cost matrix is:

$$\begin{array}{|c|c|c|} \hline & C & D \\ \hline C & 1,1 & 3,0 \\ \hline D & 0,3 & 2,2 \\ \hline \end{array} \quad (1)$$

Static game example

The prisoner's dilemma

Two possible actions: $\{C, D\}$.
The cost matrix is:

$$\begin{array}{|c|c|c|} \hline & C & D \\ \hline C & 1,1 & 3,0 \\ \hline D & 0,3 & 2,2 \\ \hline \end{array} \quad (1)$$

Lemma

There exists a unique Nash equilibrium that consists in playing D.

Do the equilibria converge?

$$\begin{tikzcd}[column sep=huge] \text{Static game } (N \text{ players}) \arrow[r, "\text{repeat}"] \arrow[d, "N \to \infty"] & \text{Stochastic (repeated games)} \\ \text{population games} \arrow[r, "\text{repeat}"] & \text{Mean-field games} \end{tikzcd}$$

Do the equilibria converge?

$$
\begin{tikzcd}
\text{Static game (N players)} \arrow[r, "\text{repeat}"] \arrow[d, "N \to \infty"] & \text{Stochastic (repeated games)} \arrow[d, "??N \to +\infty??"] \\
\text{population games} \arrow[r, "\text{repeat}"] & \text{Mean-field games}
\end{tikzcd}
$$

Stochastic Games with Identical Players

Introduced by Shapley, 1953.

Here, players are interchangeable: the dynamics, the costs and the strategies only depend on the *population distribution*.

State at time t:

$\mathbf{X}(t) = (X_1(t), \ldots, X_n(t), \ldots, X_N(t))$, with $X_n(t) \in \mathcal{S}$ (finite set).

Stochastic Games with Identical Players

Introduced by Shapley, 1953.

Here, players are interchangeable: the dynamics, the costs and the strategies only depend on the *population distribution*.

State at time t:

$\mathbf{X}(t) = (X_1(t), \ldots, X_n(t), \ldots, X_N(t))$, with $X_n(t) \in \mathcal{S}$ (finite set).

evolves in continuous time: player n takes actions $A_n(t) \in \mathcal{A}$ at instants distributed w.r.t. a Poisson process, independently of the others.

Players interact according to a mean-field model:

$$\mathbf{P}\left[X_n(t+dt)=j \middle| X_n(t)=i, A_n(t)=a, \mathbf{M}(t)=\mathbf{m}\right] = P_{ij}(a,\mathbf{m})dt$$

Strategy of a player: $\pi : (X(t), m) \mapsto A(t)$.

Players interact according to a mean-field model:

$$\mathbf{P}\left[X_n(t+dt)=j \middle| X_n(t)=i, A_n(t)=a, \mathbf{M}(t)=\mathbf{m}\right] = P_{ij}(a,\mathbf{m})dt$$

Strategy of a player: $\pi : (X(t), m) \mapsto A(t)$.

Instantaneous cost: $C(X_n(t), A_n(t), \mathbf{M}(t))$.

Player n chooses a strategy π^n to minimize her expected β-discounted payoff $V(\pi^n, \pi)$, knowing the strategies of the others:

$$V^N(\pi^n,\pi) = \mathbb{E}\left[\int e^{-\beta t} C(X_n(t), A_n(t), \mathbf{M}(t)) \middle| \begin{array}{l} A_n \text{ has d.b. } \pi^n \\ A_{n'} \text{ has d.b. } \pi \ (n' \neq \end{array}\right.$$

Definition (Nash Equilibrium)

For a given set of strategies Π, a strategy $\pi \in \Pi$ is called a symmetric Nash equilibrium in Π for the N-player game if, for any strategy $\pi^n \in \Pi$,

$$V^N(\pi, \pi) \leq V^N(\pi^n, \pi).$$

Existence is guaranteed when the dynamics and the costs are continuous functions of the population (Fink, 1964).

Mean-Field Game Model

In the mean-field limit, the population distribution $\mathbf{m}^{\pi}(t) \in \mathcal{P}(\mathcal{S})$ satisfies the mean-field equation:

$$\dot{m}_j^{\pi}(t) = \sum_{i \in \mathcal{S}} \sum_{a \in \mathcal{A}} m_i^{\pi}(t) Q_{ij}(a, \mathbf{m}^{\pi}(t)) \pi_{i,a}(\mathbf{m}^{\pi}(t)). \tag{2}$$

Mean-Field Game Model

In the mean-field limit, the population distribution $\mathbf{m}^\pi(t) \in \mathcal{P}(\mathcal{S})$ satisfies the mean-field equation:

$$\dot{m}_j^\pi(t) = \sum_{i \in \mathcal{S}} \sum_{a \in \mathcal{A}} m_i^\pi(t) Q_{ij}(a, \mathbf{m}^\pi(t)) \pi_{i,a}(\mathbf{m}^\pi(t)). \qquad (2)$$

We focus on a particular player, that we call Player 0.
Thanks to the decoupling assumption, the $P(X_0 = j) = x_j$ satisfies:

$$\dot{x}_j(t) = \sum_{i \in \mathcal{S}} \sum_{a \in \mathcal{A}} x_i(t) Q_{ij}(a, \mathbf{m}^\pi(t)) \pi^n_{i,a}(t). \qquad (3)$$

The discounted cost of Player 0 is

$$V(\pi^0, \pi) = \int_0^\infty \left(\sum_{i \in \mathcal{S}} \sum_{a \in \mathcal{A}} x_i(t) C_{i,a}(\mathbf{m}^\pi(t)) \pi^0_{i,a}(\mathbf{m}^\pi(t)) e^{-\beta t} \right) dt,$$

Definition (Mean-Field Equilibrium)

A strategy is a (symmetric) mean-field equilibrium if

$$V(\pi^{MFE}, \pi^{MFE}) \leq V(\pi, \pi^{MFE}).$$

Convergence of continuous policies

Theorem (Existence of equilibrium, Doncel, G., Gaujal 2016)

Assume that $Q_{ij}(a, \mathbf{m})$ and $C_{ia}(\mathbf{m})$ are continuous in $\mathbf{m}$. Then, there always exists a mean-field equilibrium.

Applying the Kakutani fixed point theorem for infinite dimension spaces to the population distribution (instead of directly to strategies). Does not require convexity assumptions as in Gomes, Mohr, Souza, 2013.

Theorem (Convergence, Tembine et al., 2009)

If $C_{i,a}(\mathbf{m})$, $Q_{ij}(a, \mathbf{m})$ and the policy $\pi_i(\mathbf{m})$ are continuous in $\mathbf{m}$ then the population of the finite game converges to the solution of the differential equation (2) and the evolution of one player converges to the solution of (3).

Question: where is the catch?

Non-convergence in General

We consider a matching game version of the prisoner's dilemma. The state space: $\mathcal{S} = \{C, D\}$ and $\mathcal{A} = \mathcal{S}$. Population distribution is $\mathbf{m} = (m_C, m_D)$. Cost of a player:

$$C_{i,i}(\mathbf{m}) = \begin{cases} m_C + 3m_D & \text{if } i = C \\ 2m_D & \text{if } i = D \end{cases}$$

This is the expected cost of a player matched with another player at random and using the cost matrix:

$$\begin{array}{|c|c|c|} \hline & C & D \\ \hline C & 1,1 & 3,0 \\ \hline D & 0,3 & 2,2 \\ \hline \end{array} \tag{4}$$

Lemma

There exists a unique mean-field equilibrium π^∞ that consists in always playing D.

Non-convergence in General (II)

Let us define the following stationary strategy for N players:

$$\pi^N(\mathbf{M}) = \begin{cases} D & \text{if } M_C < 1 \\ C & \text{if } M_C = 1. \end{cases}$$

"play C as long as everyone else is playing C. Play D as soon as another player deviates to D."

Non-convergence in General (II)

Let us define the following stationary strategy for N players:

$$\pi^N(\mathbf{M}) = \begin{cases} D & \text{if } M_C < 1 \\ C & \text{if } M_C = 1. \end{cases}$$

"play C as long as everyone else is playing C. Play D as soon as another player deviates to D."

Lemma

For $\beta < 1$ and N large, π^N is a sub-game perfect equilibrium of the N-player stochastic game.

Non-convergence in General (proofs)

Assume that all players, except player 0, play the strategy π^N and let us compute the best response of player 0.

If at time t_0, $M_C < 1$, then the best response of player 0 is to play D.

Non-convergence in General (proofs)

Assume that all players, except player 0, play the strategy π^N and let us compute the best response of player 0.
If at time t_0, $M_C < 1$, then the best response of player 0 is to play D.
If $M_C = 1$ then using π, has a cost
$\frac{1}{N}\sum_{i=0}^{\infty} e^{-\beta i/N} = \int \exp(-\beta t)dt + O(1/N) = 1/\beta + O(1/N)$.
If player 0 chooses action D, all players will also play D after the next step. This implies that $M_D(t) \approx 1 - \exp(-t)$ and that the player 0 will suffer a cost equal to
$\int_0^{\infty}(x_C(t) + 2 - 2e^{-t})e^{-\beta t}dt + O(1/N) \geq 2/(\beta(\beta+1)) + O(1/N)$.

Non-convergence in General (proofs)

Assume that all players, except player 0, play the strategy π^N and let us compute the best response of player 0.
If at time t_0, $M_C < 1$, then the best response of player 0 is to play D.
If $M_C = 1$ then using π, has a cost
$\frac{1}{N}\sum_{i=0}^{\infty} e^{-\beta i/N} = \int \exp(-\beta t)dt + O(1/N) = 1/\beta + O(1/N)$.
If player 0 chooses action D, all players will also play D after the next step. This implies that $M_D(t) \approx 1 - \exp(-t)$ and that the player 0 will suffer a cost equal to
$\int_0^{\infty}(x_C(t) + 2 - 2e^{-t})e^{-\beta t}dt + O(1/N) \geq 2/(\beta(\beta+1)) + O(1/N)$.
This shows that when $\beta < 1$, player 0 has no incentive to deviate from the strategy π^N so that, π^N is a sug-game perfect equilibrium.

Mean-field Games: Conclusion

With repeated game with a finite number of players, it is possible to define many equilibria by using the *"tit for tat"* principle (Folk Theorem).

Mean-field Games: Conclusion

With repeated game with a finite number of players, it is possible to define many equilibria by using the *"tit for tat"* principle (Folk Theorem).

When the number of players is infinite, the deviation of a single player is not visible by the population, the equilibria based on the "tit for tat" principle do not scale at the mean-field limit.

- This is all the more damaging because these equilibria have very good social costs: mean-field games fail to describe the best equilibria.

Are mean-field games good models?

Outline

Recap

1. Mean-field $\approx$ decoupling assumption
   - Valid for finite time.
   - Infinite horizon should be handle with care
2. Rate of convergence
   - $O(1/\sqrt{N})$ under a Lipschitz condition.
3. Controlled problems
   - OK for centralized control
   - Not that OK for games

Thank you!

http://mescal.imag.fr/membres/nicolas.gast

nicolas.gast@inria.fr

Mean-field and decoupling

Benaïm, Le Boudec 08 — *A class of mean field interaction models for computer and communication systems*, M.Benaïm and J.Y. Le Boudec., Performance evaluation, 2008.

Le Boudec 10 — *The stationary behaviour of fluid limits of reversible processes is concentrated on stationary points.*, J.-Y. L. Boudec. , Arxiv:1009.5021, 2010

Darling Norris 08 — *R. W. R. Darling and J. R. Norris*, Differential equation approximations for Markov chains, Probability Surveys 2008

G. 16 — *Construction of Lyapunov functions via relative entropy with application to caching*, Gast, N., ACM MAMA 2016

Budhiraja et al. 15 — *Limits of relative entropies associated with weakly interacting particle systems.*, A. S. Budhiraja, P. Dupuis, M. Fischer, and K. Ramanan, , Electronic journal of probability, 20, 2015.

References (continued)

Optimal control and mean-field games:

G.,Gaujal Le Boudec 12 — *Mean field for Markov decision processes: from discrete to continuous optimization*, N.Gast,B.Gaujal,J.Y.Le Boudec, IEEE TAC, 2012

G. Gaujal 12 — *Markov chains with discontinuous drifts have differential inclusion limits.*, Gast N. and Gaujal B., Performance Evaluation, 2012

Puterman — *Markov decision processes: discrete stochastic dynamic programming*, M.L. Puterman, John Wiley & Sons, 2014.

Lasry Lions — *Mean field games*, J.-M. Lasry and P.-L. Lions, Japanese Journal of Mathematics, 2007.

Tembine at al 09 — *Mean field asymptotics of markov decision evolutionary games and teams*, H. Tembine, J.-Y. L. Boudec, R. El-Azouzi, and E. Altman., GameNets 00

Applications: caches, bikes

Don and Towsley — *An approximate analysis of the LRU and FIFO buffer replacement schemes*, A. Dan and D. Towsley., SIGMETRICS 1990

G. Van Houdt 15 — *Transient and Steady-state Regime of a Family of List-based Cache Replacement Algorithms.*, Gast, Van Houdt., ACM Sigmetrics 2015

Fricker-Gast 14 — *Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity.*, C. Fricker and N. Gast. , EJTL, 2014.

Fricket et al. 13 — *Mean field analysis for inhomogeneous bike sharing systems*, Fricker, Gast, Mohamed, Discrete Mathematics and Theoretical Computer Science DMTCS

G. et al 15 — *Probabilistic forecasts of bike-sharing systems for journey planning*, N. Gast, G. Massonnet, D. Reijsbergen, and M. Tribastone, CIKM 2015