Introduction to modern Game Theory

Vassili N. Kolokoltsov

Vega, 2021

Part 4

Games of many players in statistical limit. Mean field games.

Exposition is essentially based on:

V. N. Kolokoltsov and O. A. Malafeyev. Many Agent Games in Socio-economic Systems: Corruption, Inspection, Coalition Building, Network Growth, Security. Springer Nature, 2019.

V. N. Kolokoltsov, O. A. Malafeyev. Understanding Game Theory. World Scientific, Sec. Edition 2020.

V. Kolokoltsov, *The evolutionary game of pressure (or interference), resistance and collaboration*. MOR (Mathematics of Operations Research), 42 (2017), 915 – 944.

V. N. Kolokoltsov and A. Bensoussan. Mean-field-game model of botnet defence in cyber-security (2015). AMO (Applied Mathematics and Optimization) 74(3), 669-692.

Some fundamental references for MFG theory

J-M. Lasry, and P-L. Lions, Jeux à champ moyen. I. Le cas stationnaire (French), Comptes Rendus Mathématique 343 (9), 619 - 625 (2006).

M. Huang, R. Malhamé, and P. Caines. Large population stochastic dynamic games. Commun. Inf. Syst. 6, 221 - 252 (2006).

A. Bensoussan, J. Frehse, and Ph. Yam, Mean field games and mean field type control theory. Springer, New York, 2013.

D.A. Gomes, E.A. Pimentel and V. Voskanyan, Regularity Theory for Mean-Field Game Systems. Springer, 2016.

R. Carmona, and F. Delarue, Probabilistic Theory of Mean Field Games with Applications, v. 1, 2. Springer 2018.

Big chunks of the literature devoted to inspection, corruption, terrorism, cyber-security, etc (published often in specialized journals) discuss identically the same mathematical models, unaware of this fact.

Chapter 1

Introduction: Main ideas and applications of the statistical limit, dynamic law of large numbers (LLN).

Modeling of cyber-security, crime prevention, inspection and corruption, counterterrorism, epidemiology, preservation of environment, merging of firms or banks, forming coalitions.

Highlights

- Evolutionary game theory (pairwise games between randomly selected pairs of a large population of small players) extends to the case of games conducted under the 'pressure' of an external 'major player' (principal).

- Small players can copy more beneficial strategies observed from the neighbors (myopic behavior) including building coalitions.

- LLN is given by kinetic equations that extend the RD. Rest points of these equations – ϵ-Nash equilibria for games with finite number of players.

- Random input of agents leads to the models of evolutionary growth under pressure including emergence of cooperation to survive in an unfriendly environment.

Basic setting, I

Start modeling: many similar players each having a finite number of strategies (or types) $\{1, \cdots, d\}$.

Visual picture (example of an interpretation) to have in mind: these are the levels at which an agent can:
(i) trespass the law: (tax-payer evades taxes on a certain amount, corrupted agent takes bribes of a certain level, etc), or (turning the screw of moral evaluation)
(ii) build a defence system against an attacker, say, level of protection of a computer against a cyber-gangster, level of security on an airport against terrorists, level of resistance to an oppressive military force, level of protection against a virus (vaccination).
These examples already introduce a major player P (inspector, cyber-hacker, etc).

Basic setting, II

The state space of the group is $\mathbf{Z}_+^d$, the set of sequences of d non-negative integers $n = (n_1, ..., n_d)$, where each n_i specifies the number of players in the state i.

$N = n_1 + ... + n_d$ the total number of players (assume constant for now).

Principal P exerted a pressure on small players with the level of pressure (control parameter) b (may be vector-valued to diversify the interference).

Let $R_i(x, b)$ be the payoff of the strategy i against the action b of the player P when the overall distribution

$$x = (x_1, \cdots, x_d) = (n_1, \cdots, n_d)/N$$

of the strategies applies. Dependence on x is crucial: think about social norms!

Basic setting, III

Assuming that P has some strategy $b(x, N)$ let us consider the following Markov model of the interaction of the group.

With some rate $\varkappa/N$ any pair of agents can meet and discuss their payoffs. This discussion may result in the player with lesser payoff R_i switching to the strategy with the better payoff R_j, which may occur with probability proportional to $(R_j - R_i)$.

Remark. This is pure myopic. One can also take into account random mutation on global or local levels.

More formal description: exponential-waiting-time clocks

Limiting ODE, I

We are interested in the asymptotic behavior of the chains as $N \to \infty$.

Theorem 1 (to be proved in Chapter 2). Under weak technical assumptions, the limiting process turns out to be a deterministic one governed by the system of ODE

$$\dot{x}_j = \sum_i \varkappa x_i x_j [R_j(x, b(x)) - R_i(x, b(x))], \quad j = 1, ..., d,$$

which is the system of kinetic equations generalizing (and modifying) the usual replicator dynamics.

Limiting ODE, II

The behavior of the major player: The payoff of P playing against the group of small players is given by a function $B(x, b, N)$. *Best response principal* chooses the maximum point

$$b^*(x, N) = argmax\, B(x, b, N)$$

as her strategy at any time.
Assuming there exists the limit

$$\lim_{N \to \infty} b^*(x, N) = b^*(x),$$

the limiting ODE system becomes

$$\dot{x}_j = \sum_i \varkappa x_i x_j [R_j(x, b^*(x)) - R_i(x, b^*(x))], \quad j = 1, ..., d, \tag{1}$$

Fixed points of the dynamics

Above result suggests that eventually the evolution will settle down near some stable equilibrium points of the main ODE system above.

For a subset $I \subset \{1, \cdots, d\}$, let Ω_I be the collection of $x \in \Sigma_d$ s.t.

$$x_k = 0, k \in I, \text{ and } R_j(x, b^*(x)) = R_i(x, b^*(x)) \text{ for } i, j \notin I\}.$$

Theorem 2. A vector x with non-negative coordinates is a rest point of (1), that is, it satisfies the system of equations

$$\sum_i \varkappa x_i x_j [R_j(x, b^*(x)) - R_i(x, b^*(x))] = 0, \quad j = 1, ..., d, \quad (2)$$

if and only if $x \in \Omega_I$ for some $I \subset \{1, \cdots, d\}$.

Fixed points of the dynamics

Proof. For any I such that $x_k = 0$ for $k \in I$, system (2) reduces to the same system but with coordinates $k \notin I$. Hence it is sufficient to show the result for the empty I. In this situation, system (2) reduces to

$$\sum_i x_i [R_j(x, b^*(x)) - R_i(x, b^*(x))] = 0, \quad j = 1, ..., d. \tag{3}$$

Subtracting jth and kth equations of this system yields

$$(x_1 + \cdots + x_d)[R_j(x, b^*(x)) - R_k(x, b^*(x))] = 0,$$

and thus

$$R_j(x, b^*(x)) = R_k(x, b^*(x)),$$

as required.

Fixed points of the dynamics and Nash equilibria

Is there a game-theoretic interpretation of the fixed points? (Usual evolutionary dynamics result does not apply, because of the dependence on x.)
Consider the following game Γ_N of $N+1$ players (that was tacitly borne in mind when discussing dynamics). When the major player chooses the strategy b and each of N small players chooses the state i, the major player receives the payoff $B(x, b, N)$ and each player in the state i receives $R_i(x, b)$, $i = 1, \cdots, d$ (as above, with $x = n/N$ and $n = (n_1, \cdots, n_d)$ the realized occupation numbers of all the states). Thus a strategy profile of small players in this game can be specified either by a sequence of N numbers (expressing the choice of the state by each agent), or more succinctly, by the resulting collection of frequencies $x = n/N$.

Fixed points of the dynamics and Nash equilibria

As usual, a Nash equilibrium in Γ_N is a profile of strategies (x_N, b_N^*) such that for any player changing its choice unilaterally would not be beneficial, that is

$$b_N^* = b^*(x_N, N) = argmax\, B(x_N, b, N)$$

and for any $i, j \in \{1, \cdots, d\}$

$$R_j(x - e_i/N + e_j/N, b_N^*) \leq R_i(x, b_N^*). \qquad (4)$$

A profile is an ϵ-*Nash equilibrium* if these inequalities hold up to an additive correction term not exceeding ϵ. It turns out that the singular points of (1) describe all approximate Nash equilibria for Γ_N in the following precise sense:

Fixed points of the dynamics and Nash equilibria

Theorem 3. Let $R(x,b)$ be Lipschitz continuous in x uniformly in b. Let

$$\hat{R} = \sup_{i,b} \|R_i(.,b)\|_{Lip},$$

and, for $I \subset \{1, \cdots, d\}$, let

$$\hat{\Omega}_I = \{x \in \Omega_I : R_k(x, b^*(x)) \leq R_i(x, b^*(x)) \, \text{for} \, k \in I, i \notin I\}.$$

Then the following assertions hold.
(i) The limit points of any sequence x_N such that $(x_N, b^*(x_N, N))$ is a Nash equilibrium for Γ_N belong to $\hat{\Omega}_I$ for some I. In particular, if all x_N are internal points of Σ_d, then any limiting point belongs to $\Omega_\emptyset$.
(ii) For any I and $x \in \hat{\Omega}_I$ there exists an $2\hat{R}d/N$-Nash equilibrium $(x_N, b^*(x_N, N))$ to Γ_N such that the difference of any coordinates of x_N and x does not exceed $1/N$ in magnitude.

Proof

(i) Let us consider a sequence of Nash equilibria $(x_N, b^*(x_N, N))$ such that the coordinates of all x_N in I vanish. By (4) and the definition of $\hat{R}$,

$$|R_j(x_N, b^*(x_N, N)) - R_i(x_N, b^*(x_N, N))| \leq \frac{2}{N}\hat{R} \qquad (5)$$

for any $i, j \notin I$ and

$$R_k(x_N, b^*(x_N, N)) \leq R_i(x_N, b^*(x_N, N)) + \frac{2}{N}\hat{R}, \quad k \in I, i \notin I. \qquad (6)$$

Hence $x \in \hat{\Omega}_I$ for any limiting point (x, b).

Proof (cont.)

(ii) If $x \in \hat{\Omega}_I$ one can construct its $1/N$-rational approximation, namely a sequence $x_N \in \Sigma_d \cap \mathbf{Z}_+^d/N$ such that the difference of any coordinates of x_N and x does not exceed $1/N$ in magnitude. For any such x_N, the profile $(x_N, b^*(x_N, N))$ is an $2\hat{R}d/N$-Nash equilibrium for Γ_N.

Remark. The result can be extended to cases when R is only continuous (so that solutions to kinetic equations are not well defined).

Examples. Inspection

In the standard setting of inspection games with a possibly tax-evading inspectee

$$R_j(x,b) = r + (1 - p_j(x,b))r_j - p_j(x,b)f(r_j),$$

where r is the legal payoff of an inspectee, various r_j denote various amounts of not declared profit, $j = 1, \cdots, d$, $p_j(x,b)$ is the probability for the illegal behavior of an inspectee to be found when the inspector uses budget b for searching operation and $f(r_j)$ is the fine that the guilty inspectee has to pay when being discovered.

Examples. Corruption, I

The standard model of corruption 'with benevolent principal': The payoff of a possibly corrupted inspector (now taking the role of a small player) is

$$(1-p)(r+w)+p(w_0-f),$$

where r is the bribe an inspector asks from a firm to agree not to publicize its profit (and thus allowing her not to pay tax),
w is the wage of an inspector,
f the fine she has to pay when the corruption is discovered,
p the probability of a corrupted behavior to be discovered by the benevolent principal (say, governmental official).
Finally it is assumed that when the corrupted behavior is discovered the agent not only pays fine, but is also fired from the job and has to accept a lower level activity with the reservation wage w_0.

Examples. Corruption, II

Our strategic model:
we make r to be the strategy of an inspector with possible levels $r_1, \cdots, r_d$ (the amount of bribes she is taking) and the probability p of discovery to be dependent on the effort (say, budget b) of the principal and the overall level of corruption x, with fine depending on the level of illegal behavior.
This extension of the standard model leads to the payoff

$$R_j(x, b) = (1 - p_j(x, b))(r_j + w) + p_j(x, b)(w_0 - f(r_j)),$$

which is essentially identical to the inspection game above.

Resistance and collaboration

An alternative class of strategies – collaborations with the big player on various levels c. The creation of such possibilities can be considered as a strategic action of the major player (who can thus exert some control on the rules of the game).

In biological setting this is, for instance, the strategy of dogs joining humans in hunting their 'relatives' wolves or foxes (nicely described poetically as the talk between a dog and a fox in the famous novel 'Bambi, a Life in a wood'). Historical examples include the strategy of slaves helping their masters to terrorize and torture other slaves and by doing this gaining for themselves more beneficial conditions, as described e.g. in the classics 'Uncle's Tom cabin'.

Resistance and collaboration: examples

As a military example one can indicate the strategy of the part of the population on a territory occupied by foreign militaries that joins the local support forces for the occupants, for US troops in Irak this strategy being well discussed in Chapter 2 of the book B. B. De Mesquita (a CIA consultant on political) 'The Predictioneer's Game', 2010.

Alternatively, this is also the strategy of population helping police to fight with criminals and/or terrorists.

In the world of organized crime it is also a well known strategy to play simultaneously both resistance (committing crime) and collaboration (to collaborate with police to get rid of the competitors).

Environmental interpretation

Humanity versus nature: fishing, hunting, exploiting rein forest, using medication against viruses.

Level of resistance r_i for (large amount of) living creatures (fish, animal, viruses) reflects the interplay between the danger of being caught or killed and the necessity to enter the dangerous places in order to get better forage.

Example: Cyber-security, I

Botnet defense (for example, against the famous conflicker botnet), widely discussed in the contemporary literature, since botnets (zombie networks) are considered to pose the biggest threat to the international cyber-security.

Our approach adds to the existing models the networking aspects by allowing the defenders to communicate and eventually copy more beneficial strategies.

More concretely, our general model of inspection or corruption becomes almost directly applicable in this setting by the clever linguistic change of 'detected' to 'infected' and by considering the cybecriminal as the 'principal agent'!

Example: Cyber-security, II

Namely, let r_j denote the level of defense applied by an individual (computer owner) against botnet herders, which can be the level of antivirus programs installed or the measures envisaged to quickly report and repair a problem once detected (or possibly a multidimensional parameter reflecting several defense measures).

Let $p_j(x, b)$ denote the probability for a computer of being infected given the level of defense measures r_j, the effort level b of the herder (say, budget or time spent) and the overall distribution x of infected machines (this 'mean-field' parameter is crucial in the present setting, since infection propagates as a kind of epidemic).

Example: Cyber-security, III

Then, for a player with a strategy j, the cost of being (inevitably) involved in the conflict can be naturally estimated by the formula

$$R_j(x, b) = p_j(x, b)c + r_j,$$

where c is the cost (inevitable losses) of being infected (thus one should aim at minimizing this R_j, rather then maximizing it, as in our previous models).

Of course, one can extend the model to various classes of customers (or various classes of computers) for which values of c or r_j may vary and by taking into account more concrete mechanisms of virus spreading.

Example: Terrorist attacks and counter-terrorist measures, I

We again suggest here a natural extension of the basic models to allow for the interaction of a large number of players and of various levels of attacks.

In the literature, the counterterrorists' measures are usually decomposed into two groups: proactive (or preemptive), like direct retaliation against the state-sponsor and defensive (also referred to as deterrence), like strengthening security at an airport, with the choice between the two considered as the main strategic parameter.

It is of course natural (and was mentioned in the existing literature as a desired extension) to consider various levels of the exploitation of these two strategies (our former r_i).

Example: Terrorist attacks and counter-terrorist measures, II

The first group of action is 'characterized in the literature as a pure public good, because a weakened terrorist group poses less of a threat to all potential targets', but on the other hand, it 'may have a downside by creating more grievances in reaction to heavy-handed tactics or unintended collateral damage' (because it means to 'bomb alleged terrorist assets, hold suspects without charging them, assassinate suspected terrorists, curb civil freedoms, or impose retribution on alleged sponsors'), which may result in the increase of terrorists' recruitment.

Thus, the basic model includes the recruitment benefits of terrorists as a positively correlated function of preemption efforts.

Example: Terrorist attacks and counter-terrorist measures, III

A direct extension of the basic model in the line indicated above (large number of players and the levels of attacks) suggests to write down the reward of a terrorist, or a terrorist group, considered as a representative of a large number of small players, using one of the levels of attack $j = 1, \cdots, d$ (in literature so far they considered only two levels, normal and spectacular), to be

$$R_j(x, b) = (1 - p_j(x, b)) r_j^{fail}(b) + p_j(x, b)(S_j + r_j^{succ}(b)),$$

where $p_j(x, b)$ is the probability of a successful attack (which depends on the level b of preemptive efforts of the principal b and the total distribution of terrorists playing different strategies), S_j is the direct benefits in case of a success and $r_j^{fail}(b)$, $r_j^{succ}(b)$ are the recruitment benefits in the cases of failure or success respectively.

Example: Terrorist attacks and counter-terrorist measures, IV

It is seen directly that we are again in the same situation as in all above models (up to constants and notations).

The costs of the principal are given by

$$B(x,b) = \sum_j x_j \left[(1 - p_j(x,b))b + p_j(b)(b + S_j)\right].$$

The model extends naturally to account for possibility of the actions of two types, preemption and deterrence, and to several major players (USA and EU are often considered).

Stability Analysis (Comments), I

All examples above:

$$R_j = r + (1 - F_j(b,x))r_j - F_j(b,x)f(r_j) = r + r_j - F_j(b,x)g(r_j).$$

And the evolution equations:

$$\dot{x}_j = x_j[r_j - F_j(b,x)g(r_j) - \sum_i (r_i - F_i(b,x)g(r_i))x_i].$$

Simplifying assumption $F_i(b,x) = F(b)$, $b \in \mathbf{R}$. Then, if the principle chooses $b = b^*(x)$,

$$\dot{x}_j = x_j[r_j - F(b^*(x))g(r_j) - \mathbf{E}(r - F(b^*(x))g(r))].$$

Here F and g are increasing functions.
Note that g specifies the punishment mechanism that can be chosen by the principal in his 'mechanism design' in order to shift equilibria in the desired direction.

Stability Analysis (Comments), II

Remarkable result: if g is linear $g(r) = \lambda r$, equation turns to

$$\dot{x}_j = x_j(1 - F(b)\lambda)(r_j - \mathbf{E}(r)),$$

and equation $1 - F(b)\lambda = 0$ defines a hyperplane of internal fixed points (all strategies are played with a positive probability) that are STABLE.

If g is strictly convex or concave, any fixed point can be supported only on two strategies and all these 'unpleasant' internal stable points DISAPPEAR!

Clear PRACTICAL interpretation!

Merging, splitting, coalition building (Comments), I

If any randomly chosen pair of coalitions of sizes j and k can merge with the rates $C_{kj}(x, b)$, which may depend on the whole composition x and the control parameter b of the major player, and any randomly chosen coalition of size j can split (break, fragment) into two groups of sizes $k < j$ and $j - k$ with rate $F_{jk}(x, b)$, the limiting deterministic evolution of the state is known to be described by the system of the so-called Smoluchovski equations

$$\dot{x}_k = f_k(x) = \sum_{j<k} C_{j,k-j}(x, b) x_j x_{k-j} - 2 \sum_j C_{kj}(x, b) x_j x_k$$

$$+2 \sum_{j>k} F_{jk}(x, b) x_j - \sum_{j<k} F_{kj}(x, b) x_k.$$

We added here the dependence of the coefficients C and F on x and the control parameter b.

Merging, splitting, coalition building (Comments), II

Let $R_j(x, b)$ be the payoff for the member of a coalition of size j. In our strategic setting, the rates $C_{jk}(x, b)$ and $F_{jk}(x, b)$ should depend on the differences of these rewards before and after merging or splitting. For instance, the simplest choices can be

$$C_{kj}(x, b) = a_{j+k,k}\mathbf{1}_{R_{k+j}\geq R_k}(R_{k+j}-R_k)+a_{j+k,j}\mathbf{1}_{R_{k+j}\geq R_j}(R_{k+j}-R_j),$$

with some constants $a_{lk} \geq 0$ reflecting the assumption that merging may occur whenever it is beneficial for all members concerned but weighted according to the size of the coalitions involved. Similarly

$$F_{kj}(x, b) = \tilde{a}_{kj}\mathbf{1}_{R_j\geq R_k}(R_j - R_k) + \tilde{a}_{k,k-j}\mathbf{1}_{R_{k-j}\geq R_k}(R_{k-j} - R_k).$$

Preferential attachment (Comments)

The most studied form of preferential attachment evolves by the discrete time injections of agents: with time intervals τ a new agent enters the system. With probability $\alpha(x, b)$ (which, unlike the standard model, can now depend on the distribution x and the control parameter b of the principal) the agent forms a new coalition of size 1, and with probability $1 - \alpha(x, b)$ she joins one of the existing coalitions with the probability proportional to its size (this reflects the notion of preferential attachment).

Thus if $V(x)$ is some function on the state space $h\mathbf{Z}_+^{fin}$, its expected value after a single entry changing x to $\hat{x}$ changes to

$$\mathbf{E}V(\hat{x}) = \alpha V(x + he_1) + (1 - \alpha) \sum_{k=1}^{\infty} \frac{kn_k}{L(n)} V(x - he_k + he_{k+1}),$$

where $L(n) = \sum kn_k$.

This leads necessarily to models with an infinite (say, countable) state space.

Chapter 2

Dynamic law of large numbers: rigorous results.

Convergence of Markov chains for many players confronting a major player to a limiting deterministic evolution (kinetic equation).

Preliminaries: Markov chains

A Markov chain on $\{1, ..., d\}$ is specified by the choice of a *Q-matrix* or a *Kolmogorov's matrix* Q, which is a $d \times d$ square matrix such that its non-diagonal elements are non-negative and the elements of each row sum up to zero (and thus the diagonal elements are non-positive).

The *Markov chain* with the Q-matrix Q is the process evolving by the following rule. Starting from any time t and a current state i one waits a $|Q_{ii}|$-exponential random waiting time τ and then the position jumps to a state j according to the distribution $Q_{ij}/|Q_{ii}|$. At time $t+\tau$ the procedure starts again from position j, etc.

Markov chains: analytic description

Let $X_i(t)$ denote the position of the chain at time $t \geq 0$, if it started at i at the initial time $t = 0$, and let $P_{ij}(t)$ denote the probability of the transfer from i to j in time t, so that

$$\mathbf{P}(X_j(t) = j) = P_{ij}(t).$$

Then one can show that these *transition probabilities* $P(t) = \{P_{ij}(t)\}$ satisfies the following *Kolmogorov's forward equations*

$$\frac{d}{dt} P_{ij}(t) = \sum_{l=1}^{d} Q_{lj} P_{il}(t), \quad t \geq 0,$$

or in matrix form

$$\frac{d}{dt} P(t) = P(t) Q.$$

Markov chains: analytic properties

On the other hand, the evolution of averages,

$$T^t f(i) = \mathbf{E} f(X_i(t)) = \sum_{j=1}^{d} P_{ij}(t) f(j)$$

for any function f on the state space $\{1, ..., d\}$, i.e. for any vector $f \in \mathbf{R}^d$, satisfies the *Kolmogorov's backward equations*

$$\frac{d}{dt} T^t f(i) = \sum_{j=1}^{d} Q_{ij} T^t f(j), \quad t \geq 0,$$

or in the vector form

$$\frac{d}{dt} T^t f = Q T^t f.$$

Consequently $T^t f = e^{tQ} f$.

Remark. The values of f above will be denoted both by f_n and $f(n)$ reflecting two interpretations of f as a vector in $\mathbf{R}^d$ and a function on the state space.

Generators and transition operators

The matrix Q considered as the linear operator in $\mathbf{R}^d$ is called the *generator* of the Markov chain $X_i(t)$ and the operators T^t in $\mathbf{R}^d$ are called the *transition operators*. These operators form a semigroup, that is, they satisfy the equation $T^t T^s = T^{t+s}$ for any $s, t > 0$, which is directly seen from the exponential representation $T^t = e^{tQ}$.

Non-homogeneous Markov chains

Let $\{Q(t)\} = \{(Q_{ij})(t)\}$ be a family of $d \times d$ square Q-matrices or Kolmogorov matrices depending piecewise continuously on time $t \geq 0$. The family $\{Q(t)\}$ specifies a (time non-homogeneous) *Markov chain* $X_{s,j}(t)$ on the state space $\{1, ..., d\}$.
Let us denote $X_{s,j}(t)$ the position of this process at time t if it was initiated at time s in the state i.
The *transition probabilities* $P(s,t) = (P_{ij}(s,t))_{i,j=1}^{d}$, $s \leq t$, defining the probabilities to migrate from i to j during the time segment $[s,t]$, are said to form the *transition matrix* and the corresponding operators

$$U^{s,t}f(i) = \mathbf{E}f(X_{s,i}(t)) = \sum_{j=1}^{d} P_{ij}(s,t)f(j)$$

are called the *transition operators* of the Markov chain.

Non-homogeneous Markov chains

The matrices $Q(t)$ define the time-dependent generator:

$$(Q(t)f)_n = \sum_{m \neq n} Q_{nm}(t)(f_m - f_n), \quad f = (f_1, \cdots, f_d).$$

The transition matrices satisfy the *Kolmogorov's forward equation*

$$\frac{d}{dt} P_{ij}(s,t) = \sum_{l=1}^{d} Q_{lj}(t) P_{il}(s,t), \quad s \leq t,$$

and the transition operators of this chain satisfy the *chain rule* (or *Chapman-Kolmogorov equation*): $U^{s,r} U^{r,t} = U^{s,t}$ for $s \leq r \leq t$, and the *Kolmogorov backward equations*

$$\frac{d}{ds}(U^{s,t} f)(i) = -\sum_{j=1}^{d} Q_{ij}(s)(U^{s,t} f)(j), \quad s \leq t.$$

A two-parameter family $U^{s,t}$, $s \leq t$, satisfying the chain rule, is called a (backward) *propagator*.

Discrete-time Markov chains

Concerning discrete time Markov chains $X_i(t)$ specified by the transition matrix $P = \{P_{ij}\}$ let us recall the evident fact that the probabilities P^n_{ij} of transitions $i \to j$ in time n (i.e. the probabilities of being in j at time t conditioned on the initial state i at time 0) form the matrix $P^n = \{P^n_{ij}\}$, which is the power of the transition matrix P: $P^n = (P)^n$, $n = 0, 1, \cdots$. The corresponding transition operators describing the dynamics of the averages act as the multiplication by P^n, and hence will be also denoted by P^n:

$$P^n f(i) = \mathbf{E} f(X_i(n)) = (P^n f)(i) = \sum_j P^n_{ij} f(j).$$

Discrete-time Markov chains

A trivial but important modification to be mentioned is the similar setting but with time between jumps being any fixed time τ.
This modification allows one to establish the close link between discrete and continuous time modeling.
Namely, for a continuous time Markov chain with the transition operators T^t specified by the Q-matrix Q, let

$$\tau < (\max_i |Q_{ii}|)^{-1}.$$

Then one can define the discrete-time Markov chain on the same state space with the transition matrix $P^\tau = \{P^\tau_{ij}\}$, where

$$P^\tau_{ii} = 1 - \tau |Q_{ii}|, \quad P^\tau_{ij} = \tau Q_{ij}, \quad j \neq i. \tag{7}$$

Discrete-time Markov chains

In matrix form this rewrites as

$$P^\tau = \mathbf{1} + \tau Q,$$

where $\mathbf{1}$ denotes the unit matrix. Hence

$$(P^\tau)^n = (\mathbf{1} + \tau Q)^n.$$

Consequently, if $\tau \to 0$ and $n \to \infty$ in such a way that $n\tau \to t$,

$$\lim (P^\tau)^n = e^{tQ} = T^t, \tag{8}$$

yielding the important link between the semigroup T^t and its discrete-time approximations P^τ.

Preliminaries: ODEs and first order PDEs

Let $\dot{x} = g(x)$ be a (vector-valued) ordinary differential equation (ODE) in $\mathbf{R}^d$. If g is a Lipschitz function, its solution $X_x(t)$ with the initial condition x is known to be well defined for all $t \in \mathbf{R}$. Hence one can define the lifting of this evolution on functions:

$$T^t f(x) = f(X_x(t)).$$

These *transition operators* act as continuous contraction operators on the space $C(\mathbf{R}^d)$ and in its subspace $C_\infty(\mathbf{R}^d)$, and they form a group, i.e. $T^t T^s = T^{s+t}$ for any s, t.
The operator

$$Lf(x) = g(x) \cdot \frac{\partial f}{\partial x}$$

is called the *generator* of this group, because for any $f \in C_\infty(\mathbf{R}^d) \cap C^1(\mathbf{R}^d)$

$$\frac{d}{dt}|_{t=0} T^t f = Lf.$$

Preliminaries: ODEs and first order PDEs

If $g \in C^1(\mathbf{R}^d)$, then

$$\frac{d}{dt} T^t f = L T^t f = T^t L f \tag{9}$$

for any t, so that the function $S(t, x) = T^t f(x) = f(X_x(t))$ satisfies the linear first order partial differential equation (PDE)

$$\frac{\partial S}{\partial t} - g(x) \cdot \frac{\partial S}{\partial x} = 0, \tag{10}$$

with the initial condition $S(0, x) = f(x)$ for any $f \in C_\infty(\mathbf{R}^d) \cap C^1(\mathbf{R}^d)$. Solutions $X_x(t)$ to the ODE $\dot{x} = g(x)$ are referred to as the *characteristics* of PDE (10).

ODEs: smooth dependence

Proposition.

(i) Let $g \in C_{bLip}(\mathbf{R}^d)$. Then $X_{\cdot}(t) \in C_{bLip}(\mathbf{R}^d)$ and

$$\|X_{\cdot}(t)\|_{Lip} \le \exp\{t\|g\|_{Lip}\}. \tag{11}$$

(ii) Let $g \in C^1(\mathbf{R}^d)$. Then $X_x(t) \in C^1(\mathbf{R}^d)$ as a function of x and (recall that $\|g\|_{Lip} = \|g^{(1)}\|$)

$$\left\|\frac{\partial X_x(t)}{\partial x}\right\| = \sup_{j,x}\left\|\frac{\partial X_x(t)}{\partial x_j}\right\| \le \exp\{t\|g^{(1)}\|\} = \exp\{t \sup_{k,x}\left\|\frac{\partial g(x)}{\partial x_k}\right\|\}. \tag{12}$$

Moreover, if $f \in C^1(\mathbf{R}^d)$, then $T^t f \in C^1(\mathbf{R}^d)$ and

$$\|(T^t f)^{(1)}\| = \sup_{j,x}\left|\frac{\partial}{\partial x_j} f(X_x(t))\right|$$

$$\le \|f^{(1)}\| \left\|\frac{\partial X_x(t)}{\partial x}\right\| \le \|f^{(1)}\| \exp\{t\|g^{(1)}\|\}. \tag{13}$$

ODEs: smooth dependence

Proposition (cont.)
(iii) Let $g \in C^2(\mathbf{R}^d)$. Then $X_\cdot(t) \in C^2(\mathbf{R}^d)$ and

$$\left\| \frac{\partial^2 X_x(t)}{\partial x^2} \right\| = \sup_{j,i,x} \left\| \frac{\partial^2 X_x(t)}{\partial x_i \partial x_j} \right\| \le t\|g^{(2)}\| \exp\{3t\|g^{(1)}\|\}. \quad (14)$$

Moreover, if $f \in C^2(\mathbf{R}^d)$, then $T^t f \in C^2(\mathbf{R}^d)$ and

$$\|(T^t f)^{(2)}\| = \sup_{j,i,x} \left| \frac{\partial^2}{\partial x_j \partial x_i} f(X_x(t)) \right|$$

$$\le \|f^{(2)}\| \exp\{2t\|g^{(1)}\|\} + t\|f^{(1)}\| \, \|g^{(2)}\| \exp\{3t\|g^{(1)}\|\}. \quad (15)$$

ODEs: smooth dependence

Proposition. Let $g_i \in C_{bLip}(\mathbf{R}^d)$, $i = 1, 2$ and let $X_x^i(t)$ denote the solutions of the equations $\dot{x} = g_i(x)$. Let $\|g_1 - g_2\| \le \delta$. Then

$$|X_x^1(t) - X_y^2(t)| \le (t\delta + |x - y|) \exp\{t\|g_i\|_{Lip}\}, \quad i = 1, 2. \tag{16}$$

ODEs: smooth dependence

As follows from (9),

$$\frac{T^t f - f}{t} \to Lf, \quad t \to 0,$$

for any $f \in C_\infty(\mathbf{R}^d) \cap C^1(\mathbf{R}^d)$. As a corollary to previous results let us obtain the rates of this convergence.

Proposition. Let $g \in C^1(\mathbf{R}^d)$ and $f \in C_\infty(\mathbf{R}^d) \cap C^2(\mathbf{R}^d)$. Then

$$\|T^t f - f\| \le t\|g\| \, \|f\|_{bLip}, \tag{17}$$

$$\begin{aligned} \left\|\frac{T^t f - f}{t} - Lf\right\| &\le t\|g\|(\|g\| \, \|f^{(2)}\| + \|g\|_{Lip}\|f\|_{Lip}) \\ &\le t\|g\| \, \|g\|_{bLip}\|f\|_{C^2}. \end{aligned} \tag{18}$$

Proof

By (9),

$$\|T^t f - f\| = \|\int_0^t T^s Lf\, ds\| \leq t\|Lf\|,$$

implying (17). Next, by (9),

$$\frac{T^t f - f}{t} - Lf = \frac{1}{t}\int_0^t T^s Lf\, ds - Lf = \frac{1}{t}\int_0^t (T^s Lf - Lf)\, ds.$$

Applying (17) to the function Lf yields the first inequality in (18), the second inequality being a direct consequence.

ODEs: smooth dependence

A more general setting of non-autonomous equations $\dot{x} = g(t, x)$ is required, where g is Lipschitz in x and piecewise continuous in t. Here the solutions $X_{s,x}(t)$ with the initial point x at time s are well defined and the transition operators form a two-parameter family:

$$U^{s,t} f(x) = f(X_{s,x}(t)). \tag{19}$$

These operators satisfy the chain rule $U^{s,r} U^{r,t} = U^{s,t}$ and the function $S(s, x) = U^{s,t} f(x)$ satisfies the PDE

$$\frac{\partial S}{\partial s} + g(s, x) \cdot \frac{\partial S}{\partial x} = 0, \tag{20}$$

with the initial (or terminal) condition $S(t, x) = f(x)$.

ODEs: smooth dependence

As mentioned above, the operators $U^{s,t}$ satisfying the chain rule are said to form a *propagator*. Moreover, these operators form a *Feller propagator* meaning that they also act in the space $C_\infty(\mathbf{R}^d)$ and depend strongly continuous on s and t. That is, if $f \in C_\infty(\mathbf{R}^d)$, then $U^{s,t}f(x)$ is a continuous function $t \mapsto U^{s,t} \in C_\infty(\mathbf{R}^d)$ for any s and a continuous function $s \mapsto U^{s,t} \in C_\infty(\mathbf{R}^d)$ for any t.

ODEs: smooth dependence

The solutions $X_{s,x}(t)$ to the ODE $\dot{x} = g(t,x)$ are called the *characteristics of PDE* (20). The results and estimates above hold for the equation $\dot{x} = g(t,x)$, if all norms of functions of t,x are understood as $\sup_t$ of their norms as functions of x, say

$$\|g\|_{Lip} = \sup_t \|g(t,.)\|_{Lip}, \quad \|g^{(2)}\| = \sup_t \|g^{(2)}(t,.)\|. \quad (21)$$

ODEs: smooth dependence

With this agreement the following estimates hold:

$$\|X_{s,x}(t)\|_{Lip} \leq \exp\{(t-s)\|g\|_{Lip}\}, \tag{22}$$

$$\|(U^{s,t}f)^{(1)}\| = \sup_{j,x}\left|\frac{\partial}{\partial x_j} f(X_{s,x}(t))\right| \leq \|f^{(1)}\| \exp\{(t-s)\|g^{(1)}\|\}, \tag{23}$$

$$\left\|\frac{\partial^2 X_{s,x}(t)}{\partial x^2}\right\| \leq (t-s)\|g^{(2)}\| \exp\{3(t-s)\|g^{(1)}\|\}, \tag{24}$$

$$\begin{aligned} \|(U^{s,t}f)^{(2)}\| &\leq \|f^{(2)}\| \exp\{2(t-s)\|g^{(1)}\|\} \\ &+ (t-s)\|f^{(1)}\| \, \|g^{(2)}\| \exp\{3(t-s)\|g^{(1)}\|\}, \end{aligned} \tag{25}$$

with the derivatives on l.h.s. existing whenever the derivatives on the r.h.s. exist.

ODEs: smooth dependence

As is easily seen, for propagators estimate (17) still holds:

$$\|U^{s,t}f - f\| \leq (t-s)\|g\| \, \|f\|_{bLip},$$

and in estimate (18) an additional term appears on the r.h.s. depending on the continuity of g with respect to t. For instance, if g is Lipschitz in t with the constant

$$\varkappa = \sup_x \|g(.,x)\|_{Lip},$$

the analog of estimate (18) writes down as

$$\|\frac{U^{s,t}f - f}{t-s} - L_s f\| \leq (t-s)\|g\| \, \|g\|_{bLip}\|f\|_{C^2} + (t-s)\varkappa\|f\|_{bLip}. \tag{26}$$

Remark. Evolution $(s,x) \to X_{s,x}(t)$ can be looked at as the deterministic Markov process with the transition operators $U^{s,t}$.

Linear ODEs: sensitivity

Finally let us remind the result on the sensitivity of linear evolutions with respect to a parameter. Let us consider the following linear ODE in $\mathbf{R}^d$:

$$\dot{y} = A(t,x)y, \quad y \in \mathbf{R}^d, \quad t \in [0,T], \tag{27}$$

with the matrix A depending smoothly on a parameter $x = (x_1, \cdots, x_n) \in \mathbf{R}^n$. Here we shall work with y using their sup-norms. All function norms of A will be meant as $\sup_{t\in[0,T]}$ of the corresponding norms as functions of x. For instance,

$$\|A\| = \sup_{t\in[0,T],x} \sup_i \sum_j |A_{ij}(t,x)|.$$

For the solution $y(t) = y(t,x)$ to (27) with some initial data $y_0(x)$ that may itself depend on x we have the evident estimate

$$\|y(t)\|_{sup} \le \exp\{t\|A\|\}\|y_0\|_{sup}.$$

Linear ODEs: sensitivity

Proposition. Let $A(t,.) \in C^2(\mathbf{R}^n)$. Then $y(t)$ is also twice differentiable with respect to x and

$$\left\|\frac{\partial y}{\partial x_i}\right\|_{sup} \le \exp\{2t\|A\|\}\left(\left\|\frac{\partial y_0}{\partial x_i}\right\|_{sup} + t\left\|\frac{\partial A}{\partial x_i}\right\| \|y_0\|_{sup}\right), \tag{28}$$

$$\left\|\frac{\partial^2 y}{\partial x_i \partial x_j}\right\|_{sup} \le \exp\{3t\|A\|\}\left(\left\|\frac{\partial^2 y_0}{\partial x_i \partial x_j}\right\|_{sup} + t\left\|\frac{\partial^2 A}{\partial x_i \partial x_j}\right\| \|y_0\|_{sup}\right.$$
$$\left. +2t\sup_k\left\|\frac{\partial y_0}{\partial x_k}\right\|_{sup}\sup_k\left\|\frac{\partial A}{\partial x_k}\right\|_{sup} + 2t^2\sup_k\left\|\frac{\partial A}{\partial x_k}\right\|^2\|y_0\|_{sup}\right). \tag{29}$$

Proof

All these bounds arise from the solutions of the linear equations satisfied by the derivatives in question:

$$\frac{d}{dt}\frac{\partial y}{\partial x_i} = A(t,x)\frac{\partial y}{\partial x_i} + \frac{\partial A}{\partial x_i}y,$$

$$\frac{d}{dt}\frac{\partial^2 y}{\partial x_i \partial x_j} = A(t,x)\frac{\partial^2 y}{\partial x_i \partial x_j} + \frac{\partial A}{\partial x_j}\frac{\partial y}{\partial x_i} + \frac{\partial A}{\partial x_i}\frac{\partial y}{\partial x_j} + \frac{\partial^2 A}{\partial x_i \partial x_j}y.$$

For the justification (existence of the derivatives) we again refer to the textbooks in ODEs mentioned above.

Mean-field interacting particle systems

Basic setting of mean-field interacting particle systems:
Let $\{Q(t,x)\} = \{(Q_{ij})(t,x)\}$ be a family of $d \times d$ square Q-matrices or Kolmogorov matrices depending Lipschitz continuously on a vector x from the closed simplex

$$\Sigma_d = \{x = (x_1, ..., x_d) \in \mathbf{R}_+^d : \sum_{j=1}^d x_j = 1\},$$

and piecewise continuously on time $t \geq 0$, so that $\|Q\|_{bLip} = \|Q\| + \|Q\|_{Lip}$, where

$$\|Q\| = \sup_{i,t,x} \sum_j |Q_{ij}(t,x)| < \infty,$$

$$\|Q\|_{Lip} = \sup_{x \neq y} \sup_{i,t} \frac{\sum_j |Q_{ij}(t,x) - Q_{ij}(t,y)|}{|x-y|}. \qquad (30)$$

Mean-field interacting particle systems

In view of the properties of Q-matrices this implies

$$\|Q\| = 2 \sup_{i,t,x} |Q_{ii}(t,x)| < \infty,$$

$$\|Q\|_{Lip} \leq 2 \sup_{x \neq y} \sup_{i,t} \frac{\sum_{j \neq i} |Q_{ij}(t,x) - Q_{ij}(t,y)|}{|x-y|}. \tag{31}$$

Define the norms:

$$\|Q\|_{C^1} = \|Q\| + \sup_i \sum_j \left| \sup_{k,t,x} \frac{\partial Q_{ij}}{\partial x_k} \right|,$$

$$\|Q\|_{C^2} = \|Q\|_{C^1} + \sup_i \sum_j \left| \sup_{k,l,t,x} \frac{\partial^2 Q_{ij}}{\partial x_k \partial x_l} \right|. \tag{32}$$

Mean-field interacting particle systems

As above this implies the estimates in terms of the transitions Q_{ij} with $j \neq i$:

$$\|Q\|_{C^1} \leq \|Q\| + 2 \sup_i \sum_{j \neq i} \left| \sup_{k,t,x} \frac{\partial Q_{ij}}{\partial x_k} \right|,$$

$$\|Q\|_{C^2} \leq \|Q\|_{C^1} + 2 \sup_i \sum_{j \neq i} \left| \sup_{k,l,t,x} \frac{\partial^2 Q_{ij}}{\partial x_k \partial x_l} \right|. \tag{33}$$

In what follows the matrices Q will depend on the additional parameter controlled by the principal, but for the moment this dependence is not relevant and will be ignored.

Mean-field interacting particle systems

Suppose we have a large number of particles distributed arbitrary among the types $\{1, ..., d\}$. More precisely our state space is $\mathbf{Z}_+^d$, the set of sequences of d non-negative integers $n = (n_1, ..., n_d)$, where each n_i specifies the number of particles in the state i. Let N denote the total number of particles in state n: $N = n_1 + ... + n_d$. For $i \neq j$ and a state n with $n_i > 0$ denote by n^{ij} the state obtained from n by removing one particle of type i and adding a particle of type j, that is n_i and n_j are changed to $n_i - 1$ and $n_j + 1$ respectively. The *mean-field interacting particle system* (in continuous time) specified by the family $\{Q\}$ is defined as the Markov chain on S with the generator

$$L_t^N f(n) = \sum_{i,j=1}^{d} n_i Q_{ij}(t, n/N)[f(n^{ij}) - f(n)]. \qquad (34)$$

Mean-field interacting particle systems

Normalizing the states to $x = n/N \in \Sigma_d \cap \mathbf{Z}_+^d/N$, leads to the generator (also denoted by L_t^N, with some abuse of notation)

$$L_t^N f(n/N) = \sum_{i=1}^{d} \sum_{j=1}^{d} n_i Q_{ij}(t, n/N)[f(n^{ij}/N) - f(n/N)], \quad (35)$$

or equivalently

$$L_t^N f(x) = \sum_{i=1}^{d} \sum_{j=1}^{d} x_i Q_{ij}(t, x) N[f(x - e_i/N + e_j/N) - f(x)], \quad (36)$$

where $x \in \mathbf{Z}_+^d/N$ and $e_1, ..., e_d$ denotes the standard basis in $\mathbf{R}^d$.

Mean-field interacting particle systems

Let us denote by $X^N(t) = X^N_{s,x}(t)$ the corresponding Markov chain and by $\mathbf{E} = \mathbf{E}_{s,x}$ the expectation with respect to this chain, where $x \in \Sigma_d \cap \mathbf{Z}^d_+/N$ denotes the initial state at time $s < t$. The transition operators of this chain will be denoted by $U_N^{s,t}$:

$$U_N^{s,t} f(x) = \mathbf{E} f(X^N_{s,x}(t)) = \mathbf{E}_{s,x} f(X^N(t)), \quad s \le t. \qquad (37)$$

Mean-field interacting particle systems

Two versions of the notations in this formula (with (s, x) attached either to $\mathbf{E}$ or to $X^N(t)$) are both standard in the theory. As any transition operators of a Markov chain, these operators satisfy the chain rule $U_N^{s,r} U_N^{r,t} = U_N^{s,t}$ for $s \leq r \leq t$ and form a propagator.
In the pressure and resistance setting

$$Q_{ij}(x, b) = \varkappa x_j (R_j(x, b) - R_i(x, b))^+$$

and generator (36) reduces to

$$L_t^N f(x) = \sum_{i=1}^{d} \sum_{j=1}^{d} \varkappa x_i x_j (R_j(t, x) - R_i(t, x))^+$$

$$\times N[f(x - e_i/N + e_j/N) - f(x)]. \qquad (38)$$

Mean-field interacting particle systems

Recall that the dual operator $(L_t^N)^*$ to the operator L_t^N is defined from the relation

$$\sum_{x=(n/N)\in\Sigma_d} [(L_t^N f)(x)g(x) - f(x)(((L_t^N)^* g)(x)] = 0.$$

By the shift of the summation index it is straightforward to see that the dual is given by the formula: $(L_t^N)^* g(y)$

$$= \sum_{i=1}^{d} \sum_{j\neq i} \big[(y_i + 1/N) Q_{ij}(t, y + e_i/N - e_j/N) g(y + e_i/N - e_j/N) - y_i Q_{ij}(t, y) g(y)\big], \qquad (39)$$

where, for convenience, it is set that $Q_{ij}(t,x) = 0$ for $x \notin \Sigma_d$. As is known, equilibrium probabilities g, solve the equation $(L^N)^* g = 0$.

Discrete-time version

In accordance with (7), *mean-field interacting systems in discrete time* related to the above discussed mean-field interacting particle system in continuous time specified by the family $\{Q\}$ is defined as the Markov chain evolving in discrete time $t = k\tau$, $k \in \mathbf{N}$, with

$$\tau \leq (N\|Q\|)^{-1} \tag{40}$$

with the transition probabilities

$$P^{\tau,t}_{nn^{ij}} = P^{\tau,t}_{nn^{ij},N} = \tau n_i Q_{ij}(t,x), \quad i \neq j, \tag{41}$$

the probability of remaining in a given state n being $1 - \tau \sum_i n_i |Q_{ii}(t,x)|$. As it follows from (41),

$$\frac{P^{\tau,t}_N f - f}{\tau} = L^N_t f \tag{42}$$

for all N, t and τ.

Dynamic LLN: smooth coefficients

As above, the functional norms of functions of two variables (t, x) with and $x \in \Sigma_d$ will mean the $\sup_t$ of their respective norms as functions of x.

Observe that, for $f \in C^1(\Sigma_d)$,

$$\lim_{N\to\infty,\, n/N\to x} N[f(n^{ij}/N) - f(n/N)] = \frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_i}(x),$$

so that

$$\lim_{N\to\infty,\, n/N\to x} L_t^N f(n/N) = \Lambda_t f(x),$$

where

$$\Lambda_t f(x) = \sum_{i=1}^{d} \sum_{j\neq i} x_i Q_{ij}(t,x)[\frac{\partial f}{\partial x_j} - \frac{\partial f}{\partial x_i}](x)$$

$$= \sum_{k=1}^{d} \sum_{i\neq k} [x_i Q_{ik}(t,x) - x_k Q_{ki}(t,x)] \frac{\partial f}{\partial x_k}(x). \qquad (43)$$

Dynamic LLN: smooth coefficients

More precisely, if $f \in C^2(\Sigma_d)$, then, by the Taylor formula,

$$L_t^N f(x) - \Lambda_t f(x) = \frac{1}{2N} \sum_{i=1}^{d} \sum_{j \neq i} x_i Q_{ij}(t,x) [\frac{\partial^2 f}{\partial x_i^2} + \frac{\partial^2 f}{\partial x_j^2} - \frac{\partial^2 f}{\partial x_i \partial x_j}](\theta),$$

for $x = n/N$, with some $\theta \in \Sigma_d$, and thus

$$\|L_t^N f - \Lambda_t f\| \leq \frac{1}{N} \|f^{(2)}\| \, \|Q\|. \qquad (44)$$

Dynamic LLN: smooth coefficients

The limiting operator $\Lambda_t f$ is a first-order PDO with characteristics solving the equations

$$\dot{x}_k = \sum_{i \neq k} [x_i Q_{ik}(t,x) - x_k Q_{ki}(t,x)] = \sum_{i=1}^{d} x_i Q_{ik}(t,x), \quad k = 1, ..., d, \tag{45}$$

called the *kinetic equations* for the process of interaction described above. In vector form this system rewrites as

$$\dot{x} = Q^T(t,x)x = xQ(t,x), \tag{46}$$

where Q^T is the transpose matrix to Q.
The corresponding transition operators act on $C(\Sigma_d)$ as

$$U^{s,t} f(x) = f(X_{s,x}(t)), \quad s \leq t. \tag{47}$$

Dynamic LLN: smooth coefficients

For the case of operator (38) the limiting operator takes the form

$$\Lambda_t f(x) = \sum_{i,j=1}^{d} \varkappa x_i x_j [R_j(t,x) - R_i(t,x)]^+ \left[\frac{\partial f}{\partial x_j} - \frac{\partial f}{\partial x_j}\right]$$

$$= \sum_{i,j=1}^{d} \varkappa x_i x_j [R_j(t,x) - R_i(t,x)] \frac{\partial f}{\partial x_j}, \qquad (48)$$

and the characteristics (or the kinetic equations) become

$$\dot{x}_j = \sum_{i=1}^{d} \varkappa x_i x_j [R_j(t,x) - R_i(t,x)]. \qquad (49)$$

Dynamic LLN: smooth coefficients

For time-homogeneous case, we denote by $X_x(t)$ the characteristics and by $X_x^N(t)$ the Markov chain starting at x at time $t = 0$. The generators of these processes also become time-homogeneous: $L_t^N = L^N$, $\Lambda_t = \Lambda$.

The corresponding transition operators $U^{s,t}$ depend only on the difference $t - s$ and the operators $U^t = U^{0,t}$, defined as $U^t f(x) = f(X_x(t))$, form a semigroup.

The transition operators for the Markov chain $X_x^N(t)$ are

$$U_N^t f(x) = \mathbf{E} f(X_x^N(t)) = \mathbf{E}_x f(X^N(t)).$$

Dynamic LLN: smooth coefficients

Let us write down explicitly the straightforward estimates of the norms of the r.h.s. $g(t,x) = xQ(t,x)$ of (46), as a function of x, in terms of the norms of Q introduced in (31) and (33): $\|g\| \le \|Q\|$ and

$$\|g\|_{Lip} \le \|Q\|_{bLip}, \quad \|g\|_{bLip} \le 2\|Q\|_{bLip}, \quad \|g^{(2)}\| \le 2\|Q\|_{C^2}, \tag{50}$$

because

$$\frac{\partial g_k}{\partial x_l} = Q_{lk} + \sum_j x_j \frac{\partial Q_{jk}}{\partial x_l}, \quad \frac{\partial^2 g_k}{\partial x_l \partial x_m} = \frac{\partial Q_{lk}}{\partial x_m} + \frac{\partial Q_{mk}}{\partial x_l} + \sum_j x_j \frac{\partial^2 Q_{jk}}{\partial x_l \partial x_m}.$$

Dynamic LLN: smooth coefficients

Under (31) system (46) is well-posed in Σ_d.

Remark. (i) Unlike general setting (19), where the condition $t > 0$ was not essential, here it is essential, because the preservation of the simplex Σ_d holds only in forward time. It holds, because $\dot{x}_k \geq 0$ whenever $x_k = 0$ and $x \in \Sigma_d$, which does not allow a trajectory to cross the boundary of Σ_d.
(ii) It is seen from the structure of (46) that if $x_k \neq 0$, then $(X_x(t))_k \neq 0$ for any $t \geq 0$. Hence the boundary of Σ_d is not attainable for this semigroup, but, depending on Q, it can be gluing or not.

Exercise. Show that if all elements of Q never vanish, then the points $X_x(t)$ never belong to the boundary of Σ_d for $t > 0$, even if the initial point x does so.

Dynamic LLN: smooth coefficients

Our first objective now is to show that the Markov chains $X_x^N(t)$ do in fact converge to the deterministic evolution $X_x(t)$ in the sense that the corresponding transition operators converge (the so-called weak convergence of Markov processes), the previous arguments showing only that their generators converge on sufficiently smooth functions. We are also interested in precise rates of convergence.

The next result concerns the (unrealistic) situation with optimal regularity of all objects concerned.

Dynamic LLN: smooth coefficients

Theorem 1. Let all the elements $Q_{ij}(x)$ belong to $C^2(\Sigma_d)$ and $f \in C^2(\Sigma_d)$. Then

$$\sup_{x \in \mathbf{Z}_+^d/N} |U_N^t f(x) - U^t f(x)|$$

$$\leq \frac{t\|Q\|}{N}(\|f^{(2)}\| + 2t\|f\|_{bLip}\|Q\|_{C^2}) \exp\{3t\|Q\|_{bLip}\}, \quad (51)$$

and, for any x and n/N,

$$|U_N^t f(n/N) - U^t f(x)| \leq \exp\{3t\|Q\|_{bLip}\}$$

$$\times \left[\frac{t\|Q\|}{N}(\|f^{(2)}\| + 2t\|f\|_{bLip}\|Q\|_{C^2}) + \|f\|_{bLip}\|x - n/N\|\right]. \quad (52)$$

Dynamic LLN: smooth coefficients

Theorem 1 (cont.) Finally, if the initial states n/N converge to a point $x \in \Sigma_d$, as $N \to \infty$, then

$$\sup_{0 \le t \le T} |U_N^t f(n/N) - U^t f(x)| \to 0, \quad N \to \infty, \tag{53}$$

for any T and any $f \in C(\Sigma_d)$.

Proof

To compare the semigroups, we shall use the following *comparison of propagator formula*:

$$(U^t - U^t_N)f = U_N^{t-s} U^s|_{s=0}^t f$$

$$= \int_0^t \frac{d}{ds} U_N^{t-s} U^s f \, ds = \int_0^t U_N^{t-s} (\Lambda - L^N) U^s f \, ds. \qquad (54)$$

Let us apply this equation to an $f \in C^2(\Sigma_d)$. By (15), (50),

$$\|(U^t f)^{(2)}\| \leq (\|f^{(2)}\| + 2t\|f\|_{bLip}\|Q\|_{C^2}) \exp\{3t\|Q\|_{bLip}\}.$$

Hence by (44),

$$\|(L^N - \Lambda) U^s f\| \leq \frac{\|Q\|}{N} (\|f^{(2)}\| + 2t\|f\|_{bLip}\|Q\|_{C^2}) \exp\{3t\|Q\|_{bLip}\}.$$

Consequently (51) follows from (54).

Proof (cont.)

Equation (52) follows, because, by (13) and (50),

$$|U^t f(x) - U^t f(y)| \leq \|U^t f\|_{Lip} \|x - y\|$$

$$\leq \exp\{t\|Q\|_{bLip}\}\|f\|_{Lip} \|x - y\|.$$

The last statement is obtained because of the possibility to approximate any function $f \in C(\Sigma_d)$ by smooth functions.

By the definition of the propagators U^t, U^t_N, equations (51)-(53) can be written in terms of the averages of the Markov chains X^N_x. For instance, (53) takes the form

$$\sup_{0 \leq t \leq T} |\mathbf{E} f(X^N_{n/N}(t)) - f(X_x(t))| \to 0, \quad N \to \infty. \tag{55}$$

Dynamic LLN: smooth coefficients

For time-dependent Q everything remains the same:
Theorem 2. Let all the elements $Q_{ij}(t,x)$ belong to $C^2(\Sigma_d)$ as functions of x and are piecewise continuous as functions of t. Let $f \in C^2(\Sigma_d)$. Then

$$\sup_{x \in \mathbf{Z}_+^d/N} |U_N^{s,t} f(x) - U^{s,t} f(x)|$$

$$\leq \frac{(t-s)\|Q\|}{N} (\|f^{(2)}\| + 2(t-s)\|f\|_{bLip} \|Q\|_{C^2}) \exp\{3(t-s)\|Q\|_{bLip}\}, \quad (56)$$

and the corresponding analog of (52) holds.

Dynamic LLN: Lipschitz coefficients

Let us move now to a more realistic situation where Q is only assumed to be Lipschitz, that is, (31) holds.

Theorem 3. Let the functions $Q(t,x)$ be piecewise continuous in t and belong to C_{bLip} as functions of x with $\|Q\|_{bLip} = \sup_t \|Q(t,.)\| \le \omega$ with some ω. Suppose the initial data $x(N) = n/N$ of the Markov chains $X^N_{s,x(N)}(t)$ converge to a certain x in $\mathbf{R}^d$, as $N \to \infty$. Then these Markov chains converge to the deterministic evolution $X_{s,x}(t)$, in the weak sense:

$$|\mathbf{E} f(X^N_{s,x(N)}(t)) - f(X_{s,x}(t))| \to 0, \text{ as } N \to \infty, \tag{57}$$

or in terms of the transition operators

$$|U_N^{s,t} f(x) - U^{s,t} f(x)| \to 0, \text{ as } N \to \infty, \tag{58}$$

for any $f \in C(\Sigma_d)$, the convergence being uniform in x whenever the convergence $x(N) \to x$ is uniform.

Dynamic LLN: Lipschitz coefficients

Theorem 3 (cont.) For smooth or Lipschitz f, the following rates of convergence are valid:

$$|\mathbf{E}f(X^N_{s,x(N)}(t))-f(X_{s,x(N)}(t))| \le C(t-s)\exp\{3(t-s)\|Q\|_{bLip}\}$$

$$\times\left(\frac{(t-s)^{1/2}}{N^{1/2}}\|Q\|_{bLip}(d+\|Q\|)\|f\|_{bLip}+\frac{\|Q\|}{N}\|f^{(2)}\|\right), \tag{59}$$

$$|\mathbf{E}f(X^N_{s,x(N)}(t))-f(X_{s,x(N)}(t))|$$

$$\le C\exp\{3(t-s)\|Q\|_{bLip}\}(d+\|Q\|)\|Q\|_{bLip}\frac{(t-s)^{1/2}}{N^{1/2}}\|f\|_{bLip}, \tag{60}$$

$$|f(X_{s,x(N)}(t))-f(X_{s,x}(t))| \le \exp\{(t-s)\|Q\|_{bLip}\}\|f\|_{bLip}\|x(N)-x\| \tag{61}$$

with a constant C.

Proof

To shorten the formulas let us write down a proof for time independent Q.

The Lipshitz continuity (61) of the solutions is a consequence of general results for ODEs.

Next, since any function $f \in C(\mathbf{R}^d)$ can be approximated by functions from $C^2(\mathbf{R}^d)$, the convergence (57) follows from (59) and (61).
Thus it remains to show (59) and (60).

The main idea is to approximate all Lipschitz continuous functions involved by the smooth ones.

Proof (cont.)

Namely, choosing an arbitrary mollifier χ (non-negative infinitely smooth even function on $\mathbf{R}$ with a compact support and $\int \chi(w)\,dw = 1$) and the corresponding mollifier $\phi(y) = \prod \chi(y_j)$ on $\mathbf{R}^{d-1}$, let us define, for any function V on Σ_d, its approximation

$$\Phi_\delta[V](x) = \int_{R^{d-1}} \frac{1}{\delta^{d-1}} \phi\left(\frac{y}{\delta}\right) V(x-y)\,dy$$

$$= \int_{\mathbf{R}^{d-1}} \frac{1}{\delta^{d-1}} \phi\left(\frac{x-y}{\delta}\right) V(y)\,dy.$$

Proof (cont.)

Notice that Σ_d is $(d-1)$-dimensional object, so that any V on it can be considered as a function of first $(d-1)$ coordinates of a vector $x \in \Sigma_d$ (continued to $\mathbf{R}^{d-1}$ in an arbitrary continuous way). It follows that

$$\|\Phi_\delta[V]\|_{C^1} = \|\Phi_\delta[V]\|_{bLip} \leq \|V\|_{bLip} \tag{62}$$

for any δ and

$$\begin{aligned} |\Phi_\delta[V](x) - V(x)| &\leq \int \frac{1}{\delta^{d-1}} \phi\left(\frac{y}{\delta}\right) |V(x-y) - V(x)| \, dy \\ &\leq \|V\|_{Lip} \int_{\mathbf{R}^{d-1}} \frac{1}{\delta^{d-1}} \phi\left(\frac{y}{\delta}\right) |y|_1 \, dy \\ &\leq \delta(d-1)\|V\|_{Lip} \int_{\mathbf{R}} |w|\chi(w) \, dw. \end{aligned} \tag{63}$$

Proof (cont.)

Next, the norm $\|\Phi_\delta[V]\|_{C^2}$ does not exceed the sum of the norm $\|\Phi_\delta[V]\|_{C^1}$ and the supremum of the Lipschitz constants of the functions

$$\frac{\partial}{\partial x_j}\Phi_\delta[V](x) = \int \frac{1}{\delta^d}\left(\frac{\partial}{\partial x_j}\phi\right)\left(\frac{y}{\delta}\right)V(x-y)\,dy.$$

Hence

$$\|\Phi_\delta[V]^{(2)}\| \le \|V\|_{bLip}\frac{1}{\delta}\int |\chi'(w)|\,dw,$$

$$\|\Phi_\delta[V]\|_{C^2} \le \|V\|_{bLip}\left(1+\frac{1}{\delta}\int |\chi'(w)|\,dw\right). \qquad (64)$$

Proof (cont.)

Let $U^t_{N,\delta}$ and U^t_δ denote the same transition operators as above but built from the approximating matrices

$$\Phi_\delta[Q](x) = \int \frac{1}{\delta^d}\phi\left(\frac{y}{\delta}\right) Q(x-y)\, dy.$$

Notice that $\Phi_\delta[Q](x)$ are also Q-matrices for any δ.
Similarly we denote by $L^{N,\delta}$ and Λ^δ the corresponding generators and by $X^\delta_x(t)$ the characteristics with $\Phi_\delta[Q]$ used instead of Q.
By (63) and (16),

$$\|X_t(x) - X^\delta_t(x)\| \leq C\delta t d \|Q\|_{bLip} \exp\{t\|Q\|_{bLip}\}$$

and hence

$$\begin{aligned} |U^t f(x) - U^t_\delta f(x)| &= |f(X_t(x) - f(X^\delta_t(x))| \\ &\leq C\|Q\|_{bLip}\|f\|_{bLip}\delta t d \exp\{t\|Q\|_{bLip}\}. \qquad (65) \end{aligned}$$

Proof (cont.)

Moreover, since

$$\|(L^{N,\delta} - L^N) f\| \leq \delta(d-1)\|f\|_{bLip}\|Q\|_{bLip} \int_{\mathbf{R}} |w|\chi(w)\, dw,$$

it follows by (54) applied to propagators U_N and $U_{N,\delta}$ that the same estimate (65) holds for the difference $U^t_{N,\delta} - U^t_N$:

$$\|U^t_{N,\delta} f - U^t_N f\| \leq C\|Q\|_{bLip}\|f\|_{bLip}\delta t d \exp\{t\|Q\|_{bLip}\}. \quad (66)$$

By (51) and (64),

$$\|U^t_{N,\delta} f - U^t_\delta f\| \leq \frac{t\|Q\|}{N}(\|f^{(2)}\| + \frac{Ct}{\delta}\|Q\|_{bLip}\|f\|_{bLip}). \quad (67)$$

Proof (cont.)

Therefore,

$$\|U_N^t f - U^t f\| \leq \|U_N^t f - U_{N,\delta}^t f\| + \|U_{N,\delta}^t f - U_\delta^t f\| + \|U_\delta^t f - U^t f\|$$

$$\leq Ct \exp\{3t\|Q\|_{bLip}\}$$

$$\times \left(\delta d \|Q\|_{bLip} \|f\|_{bLip} + \frac{\|Q\|}{N} \|f^{(2)}\| + \frac{t}{N\delta} \|Q\|_{bLip} \|Q\| \|f\|_{bLip} \right)$$

Thus choosing $\delta = \sqrt{t/N}$, makes the decay rate of δ and $t/(N\delta)$ equal yielding (59).

Proof (cont.)

Finally, if f is only Lipschitz, we approximate it by $\tilde{f} = \Phi_{\tilde{\delta}}[f]$, so that the second derivative of $\Phi_{\tilde{\delta}}[f]$ is bounded by $\|f\|_{bLip}/\tilde{\delta}$. By the contraction property of U_N^t and U^t,

$$\|U_N^t(f - \tilde{f})\| \leq \|f - \tilde{f}\| \leq Cd\tilde{\delta}\|f\|_{bLip},$$

$$\|U^t(f - \tilde{f})\| \leq \|f - \tilde{f}\| \leq Cd\tilde{\delta}\|f\|_{bLip}.$$

Thus the rates of convergence for f become of order

$$[d\tilde{\delta}+t\delta d\|Q\|_{bLip}+\frac{t^2}{N\delta}\|Q\|_{bLip}\|Q\|+\frac{t}{N\tilde{\delta}}]\|f\|_{bLip}\exp\{3t\|Q\|_{bLip}\}.$$

Choosing $\delta = \tilde{\delta} = \sqrt{t/N}$ yields (60).

Dynamic LLN with major players

We are mostly interested in the presence of a principal that may exert pressure on small players on the level described by the parameter b from a bounded convex subset of a Euclidean space. Mean-field interacting particle system controlled by the principal will be generated by (34) with the coefficients depending on the control parameter b of the principal:

$$L^{N,b}f(n) = \sum_{i,j=1}^{d} n_i Q_{ij}(n/N, b)[f(n^{ij}) - f(n)]. \qquad (68)$$

Best response principle

In the simplest setting, which we refer to as 'best response principal', one can imagine the principal choosing the value of b^* maximizing some current profit $B(x, b, N)$ for given x, N:

$$b^*(x, N) = argmax B(x, ., N). \tag{69}$$

If there exists a limit $b^*(x) = \lim b^*(x, N)$, the limiting evolution (45) becomes:

$$\dot{x}_k = \sum_{i=1}^{d} x_i Q_{ik}(x, b^*(x)), \quad k = 1, ..., d, \tag{70}$$

or, in particular in pressure and resistance framework, to evolution:

$$\dot{x}_j = \sum_i \varkappa x_i x_j [R_j(x, b^*(x)) - R_i(x, b^*(x))], \quad j = 1, ..., d. \tag{71}$$

Best response principle

The corresponding modification of above results is straightforward.

Theorem 4. Assume

$$|b^*(x,N) - b^*(x)| \le \epsilon(N), \qquad (72)$$

with some $\epsilon(N) \to 0$, as $N \to \infty$ and some function $b^*(x)$, and let the functions $Q(x,b)$ (or, in particular, $R_j(x,b)$ in the pressure and resistance framework), $b^*(x,N)$, $b^*(x)$ belong to C_{bLip} as a function of their variables with norms uniformly bounded by some ω. Suppose the initial data $x(N) = n/N$ of the Markov chains $X^N_{x(N)}(t)$ converge to a certain x in $\mathbf{R}^d$, as $N \to \infty$. Then these Markov chains converge to the deterministic evolution $X_x(t)$ solving (70) (or (71) respectively):

$$|\mathbf{E}f(X^N_{x(N))}(t) - f(X_x(t))| \to 0, \text{ as } N \to \infty. \qquad (73)$$

Best response principle

Theorem 4 (cont.)
Finally, for any $f \in C(\Sigma_d)$, the convergence is uniform in x whenever the convergence $x(N) \to x$ is uniform. For Lipschitz f, estimate (61) holds and (60) generalizes to

$$|\mathbf{E} f(X^N_{x(N)}(t) - f(X_{x(N)}(t))|$$

$$\leq C(\omega, t) \left(\frac{dt^{1/2}}{N^{1/2}} + t\epsilon(N) \right) \|f\|_{bLip}. \tag{74}$$

Dynamic LLN with distinguished (tagged) player

Under the assumptions above, let us assume that one distinguished agent in the group of N players deviates from the general rules moving according to the transition Q-matrix $Q^{dev}(t,x)$. Then the natural state-space for such Markov chain will be $\{1,\cdots,d\}\times\Sigma_d$, the first coordinate j denoting the position of the tagged player. Instead of (36), the generator of this Markov chain becomes

$$L_t^{N,dev}f(j,x)=\sum_k Q_{jk}^{dev}(t,x)(f(k,x)-f(j,x))$$

$$+\sum_i(x_i-\delta_i^j/N)\sum_{k\neq i}Q_{ik}(t,x)\left[f(j,x-e_i/N+e_k/N)-f(j,x)\right]. \tag{75}$$

Dynamic LLN with distinguished (tagged) player

Let $U_N^{s,t}$ denote the transition operators of this Markov chains. For smooth f and as $N \to \infty$ operators (75) converge to the operator

$$\Lambda_t^{dev} f(j,x) = \sum_k Q_{jk}^{dev}(t,x)(f(k,x) - f(j,x))$$

$$+ \sum_i x_i \sum_{k \neq i} Q_{ik}(t,x) \left[\frac{\partial f}{\partial x_k} - \frac{\partial f}{\partial x_i} \right] (j,x), \qquad (76)$$

with the rates of convergence

$$\|L_t^{N,dev} f - \Lambda_t^{dev} f\| = \sup_{j,x} |(L_t^{N,dev} - \Lambda_t^{dev}) f(j,x)| \le \frac{\|Q\|}{N} (\|f^{(2)}\| + 2\|f\|) \qquad (77)$$

where

$$\|f^{(2)}\| = \sup_{j,i,k,x} \left| \frac{\partial f}{\partial x_i \partial x_k}(j,x) \right|.$$

Dynamic LLN with distinguished (tagged) player

Operator (76) generates quite specific Markov process on $\{1, \cdots, d\} \times \Sigma_d$ (not a chain any more). Its second coordinate x evolves according to the deterministic kinetic equations $\dot{x} = Q^T(t,x)x$, independently on the random first coordinate, which, given j, x at time s, evolves according to the time-nonhomogeneous Markov chain $J^x_{s,j}(t) \in \{1, \cdots, d\}$ with the Q-matrix

$$Q^{dev}_{ij}(t) = Q_{ij}(t, X_{s,x}(t), u^{dev}_i(t)).$$

Therefore the transition operators $U^{s,t}$ of this process can be written as

$$U^{s,t} f(j,x) = \mathbf{E} f(J^x_{s,j}(t), X_{s,x}(t)). \tag{78}$$

If f does not depend on x, then

$$U^{s,t} f(j,x) = \mathbf{E} f(J^x_{s,j}(t)). \tag{79}$$

Dynamic LLN with distinguished (tagged) player

Theorem 5. Let us assume that one distinguished agent in the group of N players deviates from the general rules moving according to the transition Q-matrix $Q^{dev}(t,x)$, $t \in [0, T]$, satisfying the same regularity assumptions as Q. Let $f(j,x) = f(j)$ does not explicitly depend on x. Then

$$\|(U^{s,t} - U_N^{s,t})f\|_{sup} \leq \frac{(t-s)^{3/2}}{N^{1/2}} C(d, T, \|Q\|_{bLip}, \|Q^{dev}\|_{bLip})\|f\|_{sup}, \tag{80}$$

with a constant C depending on $d, T, \|Q\|_{bLip}, \|Q^{dev}\|_{bLip}$. For smooth Q and Q^{dev},

$$\|(U^{s,t} - U_N^{s,t})f\|_{sup} \leq \frac{(t-s)^2}{N} C(T, \|Q\|_{C^2}, \|Q^{dev}\|_{C^2})\|f\|_{sup}, \tag{81}$$

with a constant C depending on $T, \|Q\|_{C^2}, \|Q^{dev}\|_{C^2}$.

Proof

Let us start with the case of smooth Q and Q^{dev}. Using (77) and the comparison of propagators formula (54) we derive that

$$\|(U^{s,t} - U_N^{s,t})f\|_{sup} = sup_{j,x}|(U^{s,t} - U_N^{s,t})f(j,x)|$$

$$\leq (t-s) \sup_{r\in[s,t]} \|(L_t^{N,dev} - \Lambda_t^{dev})U^{r,t}f\|_{sup}$$

$$\leq \frac{t-s}{N}\|Q\| \left(\sup_{r\in[s,t]} \|(U^{r,t}f)^{(2)}\| + 2\|f\| \right). \qquad (82)$$

Thus we need to estimate

$$\|(U^{r,t}f)^{(2)}\| = \sup_{k,l,j,x} \left| \frac{\partial^2 U^{r,t} f(j,x)}{\partial x_k \partial x_l} \right|,$$

with $U^{s,t}$ given by (79) (and $f(j,x) = f(j)$).

Proof (cont.)

To deal with $U^{s,t}$ it is convenient to fix $s < T$ and x and to consider the auxiliary propagator $U^{r,t}_{[s,x]}$, $s \le r \le t \le T$, of the Markov chain $Y_{r,j}(t)$ in $\{1, \cdots, d\}$ with the Q-matrix $Q^{dev}(t, X_{s,x}(t))$, so that

$$U^{r,t}_{[s,x]} f(j) = \mathbf{E} f(Y_{r,j}(t)), \tag{83}$$

and

$$U^{s,t} f(j) = U^{s,t}_{[s,x]} f(j). \tag{84}$$

Proof (cont.)

Unlike $U^{s,t}$ acting on functions on $\{1, \cdots, d\} \times \Sigma_d$, the propagator $U_{[}s, x]^{r,t}$ is a propagator of a usual Markov chain and hence its action satisfies the ODE

$$\frac{d}{dr} U_{[s,x]}^{r,t} f(j) = [Q^{dev}(t, X_{s,x}(t)) U_{[s,x]}^{r,t} f](j)$$

$$= \sum_k Q_{jk}^{dev}(t, X_{s,x}(t))(U_{[s,x]}^{r,t} f)(k).$$

Proof (cont.)

To find the derivatives with respect to x we can use the ODE sensitivity results, yielding

$$\sup_{i,k,x}\left|\frac{\partial}{\partial x_i}U^{r,t}_{[s,x]}f(k)\right|$$

$$\leq (t-r)\sup_k |f(k)| \sup_{x,i}\left\|\frac{\partial Q^{dev}(t, X_{s,x}(t))}{\partial x_i}\right\| \exp\{2(t-r)\|Q^{dev}\|\},$$

and

$$\sup_{i,j,k}\left|\frac{\partial^2}{\partial x_i \partial x_j}U^{r,t}_{[s,x]}f(k)\right| \leq \exp\{3(t-r)\|Q^{dev}\|\}\sup_k |f(k)|$$

$$\times\Big((t-r)\sup_{i,j}\left\|\frac{\partial^2 Q^{dev}(t, X_{s,x}(t))}{\partial x_i \partial x_j}\right\|$$

$$+(t-r)^2\sup_i\left\|\frac{\partial Q^{dev}(t, X_{s,x}(t))}{\partial x_i}\right\|^2\Big).$$

Proof (cont.)

Since

$$\frac{\partial Q^{dev}(t, X_{s,x}(t))}{\partial x_i} = \frac{\partial Q^{dev}(t,y)}{\partial y}\Big|_{y=X_{s,x}(t)} \frac{\partial X_{s,x}(t))}{\partial x_i}$$

and similarly for the second derivative, we can use estimates (12) to estimate the derivatives of $X_{s,x}(t)$ and thus to obtain $\|(U^{r,t}f)^{(2)}\| \leq C$ with C depending on $T, \|Q\|_{C^2}, \|Q^{dev}\|_{C^2}$ and hence (81) follows by (82).
When Q and Q^{dev} are only Lipschitz we use the same approximations $\Phi_\delta[Q]$ and $\Phi_\delta[Q^{dev}]$ as above and get the rate of convergence of order $td\delta + t^2/(\delta N)$ yielding (80) by choosing $\delta = \sqrt{t/N}$.

Extension to $f(j, x)$ depending on x

Theorem 6. Let us assume that one distinguished agent in the group of N players deviates from the general rules moving according to the transition Q-matrix $Q^{dev}(t, x)$, $t \in [0, T]$, satisfying the same regularity assumptions as Q. Then

$$\|(U^{s,t} - U_N^{s,t})f\|_{sup}$$

$$\leq \frac{(t-s)^{1/2}}{N^{1/2}} C(d, T, \|Q\|_{bLip}, \|Q^{dev}\|_{bLip})\|f\|_{bLip}, \quad 0 \leq s \leq t \leq T, \tag{85}$$

with a constant C depending on $d, T, \|Q\|_{bLip}, \|Q^{dev}\|_{bLip}$. For smooth Q and Q^{dev} and f,

$$\|(U^{s,t} - U_N^{s,t})f\|_{sup} \leq \frac{(t-s)}{N} C(T, \|Q\|_{C^2}, \|Q^{dev}\|_{C^2})\|f\|_{C^2}, \tag{86}$$

with a constant C depending on $T, \|Q\|_{C^2}, \|Q^{dev}\|_{C^2}$.

Simplest example: two-state models

Assume the state space of small players consists of only two strategies. Then the rates are specified just by two numbers, Q_{12} and Q_{21}, and the distribution by one number $x = n/N$, the fraction of players using the first strategy.
When dealing with time independent transitions, the limiting dynamics reduces to the following single equation:

$$\dot{x} = (1 - x)Q_{21}(x) - xQ_{12}(x). \tag{87}$$

Simplest example: two-state models

Assume now that the interval $[0,1]$ is decomposed into a finite number of regions with the directions of preferred transitions alternating between them. Namely, let

$$0 = a_0 < a_1 < \cdots < a_{k-1} < a_k = 1, \quad I_k = (a_{k-1}, a_k),$$

and let $Q_{21} > 0$, $Q_{12} = 0$ in I_k with even k, and $Q_{21} = 0$, $Q_{12} > 0$ in I_k with odd k. Assume also that $Q_{12}(x)$ and $Q_{21}(x)$ are Lipschitz continuous and vanish on the boundary points $x = 0$ and $x = 1$.

Example: The minority game. It has two intervals: $I_0 = (0, 1/2)$, $I_1 = (1/2, 1)$.

The fixed points of the dynamics are the points of 'changing interests': $x = a_j$, $j = 0, \cdots, k$. The rest points a_k with even (respectively odd) k are stable (respectively unstable) for dynamics (87). If dynamics starts at $x \in I_k$ with even (resp. odd) k, the solution $X_x(t)$ of (87) will tend to the left point a_k (resp. right point a_{k+1}) of I_k, as $t \to \infty$.

Simplest example: two-state models

The same behavior can be seen to hold for the approximating systems of N agents evolving according to the corresponding Markov $X_x^N(t)$ chain with the generator

$$L_t^N f(x) = x Q_{12}(x) N[f(x - 1/N) - f(x)] \\ + (1 - x) Q_{21}(x) N[f(x + 1/N) - f(x)]. \qquad (88)$$

Proposition (simple). Under the above assumptions and for any sufficiently large N, the Markov chain $X_x^N(t)$ starting at a point $x = n/N \in (a_k, a_{k+1})$ moves to the left (resp. right) if k is even (resp. odd), reaches an $1/N$- neighborhood of a_k (resp. a_{k+1}) in finite time and remains in this neighborhood for ever after.

Exercise. Prove this statement.

Simplest example: nonlinear Hawk-Dove game

The replicator dynamics for Hawk-Dove game (Part 2) can be obtained as the dynamic LLN for a Markov model of N interacting species, which can use one of two strategies, h or d, with each pair meeting randomly and playing the game described by the Table. The payoff for each player in this game of N players, the fraction x (resp. $1-x$) of which playing h (resp. d), is $Hx + V(1-x)$ for being a hawk and $D(1-x)$ for being a dove.

If there are only a few hawks and plenty of dove to rob, the hawks get less impetus to fight between themselves meaning that H may increase with the decay of x. Similarly, if there are too many doves on the restricted resource area, they may start behaving more aggressively meaning that D may decrease with the decay of x.

Simplest example: nonlinear Hawk-Dove game

Choosing linear functions to express this dependence in the simplest form, the modified table of this *nonlinear Hawk-and-Dove game* becomes

	hawk	dove
hawk	$H + ax$	V
dove	0	$D - bx$

with some constants $a, b > 0$.

Simplest example: nonlinear Hawk-Dove game

In this setting the game of two players stops being defined independently, but the game of N players choosing one of the strategies h or d remains perfectly well defined (with the theory above fully applied) with the payoffs of hawks and doves becoming

$$\Pi_h(x) = (H + ax)x + V(1 - x), \quad \Pi_d(x) = (D - bx)(1 - x).$$

The replicator dynamics generalizes to the kinetic equation

$$\dot{x} = x(1-x)[x(H + ax + D - bx - V) - (D - bx - V)]. \quad (89)$$

Exercise. Find the rest points of this dynamics and investigate their stability.

Chapter 3

Mean-field type dynamic control with major players.

Markov decision (or control) processes for the principal on the evolutionary background of permanently varying profiles of small players, LLN limit for discrete and continuous time, infinite horizon problems with discounting, turnpike theory.

Multi-step decision making of the principal

As above, we shall work with the case of a finite-state-space of small players, so that the state space of the group is given by vectors $x = (n_1, \cdots, n_d)/N$ from the lattice $\mathbf{Z}_+^d/N$.

Starting with a discrete time case, we denote by $X_N(t, x, b)$ the Markov chain generated by the operator

$$L^{b,N} f(x) = N \sum_i x_i Q_{ij}(x, b) \left[f\left(x - \frac{e_i}{N} + \frac{e_j}{N} \right) - f(x) \right], \tag{90}$$

and starting in $x \in \mathbf{Z}_+^d/N$ at the initial time $t = 0$.

Multi-step decision making of the principal

Let us assume that the principal is updating her strategy in discrete times $\{k\tau\}$, $k = 0, 1, \cdots ..., n-1$, with some fixed $\tau > 0$, $n \in \mathbf{N}$, aiming at finding a strategy $\pi = \{b_0, b_1, \cdots, b_{n-1}\}$ maximizing the reward

$$V_n^{\pi,N}(x(N)) = \mathbf{E}_{N,x(N)} \left[\tau B(x_0, b_0) + \cdots + \tau B(x_{n-1}, b_{n-1}) + V(x_n)\right], \tag{91}$$

where B and V are given functions (the running and the terminal payoff), $x(N) \in \mathbf{Z}_+^d/N$ also given,

$$x_k = X_N(\tau, x_{k-1}, b_{k-1}), \quad k = 1, 2, \cdots,$$

and $b_k = b_k(x_k)$ are measurable functions of the current state $x = x_k$ ($\mathbf{E}_{N,x(N)}$ denotes the expectation specified by such process).

Multi-step decision making of the principal

By the basic dynamic programming the maximal rewards $V_n^N(x(N)) = \sup_\pi V_n^{\pi,N}(x(N))$ at different times k are linked by the optimality equation $V_k^N = S[N]V_{k-1}^N$, where the Shapley (or Bellman) operator $S[N]$ is

$$S[N]V(x) = \sup_b \left[\tau B(x,b) + \mathbf{E} V(X_N(\tau, x, b))\right], \qquad (92)$$

so that

$$V_n^N = S[N]V_{n-1}^N = S^n[N]V. \qquad (93)$$

Multi-step decision making of the principal

We are interested in the LLN limit $N \to \infty$, where we expect the limiting problem for the principal to be the maximization of the reward

$$V_n^\pi(x_0) = \tau B(x_0, b_0) + \cdots + \tau B(x_{n-1}, b_{n-1}) + V_0(x_n), \quad (94)$$

where we assume the existence of

$$x_0 = \lim_{\mathbf{N}\to\infty} x(N). \quad (95)$$

Multi-step decision making of the principal

Here

$$x_k = X(\tau, x_{k-1}, b_{k-1}), \quad k = 1, 2, \cdots, \tag{96}$$

with $X(t, x, b)$ denoting the solution to the characteristic system (or kinetic equations)

$$\dot{x}_k = \sum_{i=1}^{d} x_i Q_{ik}(x, b(x)), \quad k = 1, ..., d, \tag{97}$$

with the initial condition x at time $t = 0$.

In the pressure and resistance framework,

$$\dot{x}_j = \sum_i \varkappa x_i x_j [R_j(x, b) - R_i(x, b)], \quad j = 1, ..., d. \tag{98}$$

Multi-step decision making of the principal

Again by dynamic programming, the maximal reward in this problem $V_n(x) = \sup_\pi V_n^\pi(x)$, $\pi = \{b_k\}$, is obtained by the iterations of the corresponding Shapley operator, $V_n = S^n V_0$, with

$$SV(x) = \sup_b \left[\tau B(x, b) + V(X(\tau, x, b))\right]. \qquad (99)$$

Especially for the application to the continuous time models it is important to have estimates of convergence uniform in $n = t/\tau$ for bounded total time $t = n\tau$.

Multi-step decision making of the principal

As a preliminary step let us prove a rather standard fact about the *propagation of continuity* by the operator S.

Proposition. Let V, B, Q be bounded continuous functions, which are Lipschitz continuous in x, with $\varkappa_V = \|V\|_{Lip}$ and

$$\varkappa_B = \sup_b \|B(.,b)\|_{Lip} < \infty, \quad \omega = \sup_b \|Q(.,b)\|_{bLip} < \infty. \tag{100}$$

Then $S^n V \in C_{bLip}$ for all n and

$$\|S^n V\| \leq t\|B\| + \|V\|, \quad \|S^n V\|_{Lip} \leq (t\varkappa_B + \varkappa_V)e^{t\omega} \tag{101}$$

for $t = n\tau$.

Proof

First equation in (101) follows from the definition of SV. Next, by (50) and (16),

$$|X(\tau, x, b) - X(\tau, y, b)| \leq |x - y|e^{t\omega},$$

and therefore $|SV(x) - SV(y)|$ does not exceed

$$\sup_b [\tau B(x, b) + V(X(\tau, x, b)) - \tau B(y, b) - V(X(\tau, y, b))]$$

$$\leq |x - y|(\tau \varkappa_B + \varkappa_V e^{t\omega}).$$

Proof (cont.)

Similarly,

$$|S^2V(x) - S^2V(y)|$$

$$\leq \sup_b [\tau B(x,b) + SV(X(\tau,x,b)) - \tau B(y,b) - SV(X(\tau,y,b))]$$

$$\leq |x-y|(\tau\varkappa_B + \|SV\|_{Lip} e^{t\omega})$$

$$\leq |x-y|(\tau\varkappa_B + (\tau\varkappa_B + \varkappa_V e^{t\omega})e^{t\omega}).$$

And by induction:

$$\|S^nV\|_{Lip} \leq \tau\varkappa_B(1 + e^{\tau\omega} + \cdots + e^{\tau\omega(n-1)}) + \varkappa_V e^{\tau\omega n}$$

$$\leq n\tau\varkappa_B e^{\tau\omega n} + \varkappa_V e^{\tau\omega n},$$

implying the second estimate in (101).

Theorem 1 (Multi-step nonlinear Markov control)

(i) Assume (100) and (95) hold. Then, for any $\tau \in (0,1]$, $n \in N$ and $t = \tau n$,

$$\|S^n[N]V - S^nV\| \le C(d,\omega)(t\varkappa_B + \varkappa_V)e^{t\omega}(t\sqrt{1/(\tau N)} + |x(N) - x|) \tag{102}$$

with a constant $C(d,\omega)$. In particular, for $\tau = N^{-\epsilon}$, $\epsilon \in (0,1)$,

$$\|S^n[N]V - S^nV\| \le C(d,\omega)(t\varkappa_B + \varkappa_V)e^{t\omega}(tN^{-(1-\epsilon)/2} + |x(N) - x|). \tag{103}$$

(ii) If there exists a Lipshitz continuous optimal policy $\pi = \{b_k\}$, $k = 1, \cdots, n$, for the limiting optimization problem, then π is approximately optimal for the N-agent problem, in the sense that for any $\epsilon > 0$ there exists N_0 such that, for all $N > N_0$,

$$|V_n^N(x(N)) - V_n^{N,\pi}(x(N))| \le \epsilon.$$

Proof

(i) Let $L = \sup_{k \leq n} \|S^n V\|_{Lip}$. By (101) it is bounded by $(t\varkappa_B + \varkappa_V)e^{t\omega}$. By (60),

$$|S[N]V(x) - SV(x)| \leq sup_b |\mathbf{E}V(X_x^N(\tau)) - V(X_x(\tau))|$$

$$\leq C(d,\omega)L\sqrt{\tau/N},$$

where we used $\tau \leq 1$ to estimate $e^{3\tau\omega} \leq C(\omega)$. Next,

$$|S^2[N]V(x) - S^2V(x)| \leq sup_b |\mathbf{E}S[N]V(X_x^N(\tau)) - SV(X_x(\tau))|$$

$$\leq sup_b \mathbf{E}|S[N]V(X_x^N(\tau)) - SV(X_x^N(\tau))|$$
$$+ sup_b |\mathbf{E}SV(X_x^N(\tau)) - SV(X_x(\tau))|$$
$$\leq C(d,\omega)L\sqrt{\tau/N} + C(d,\omega)L\sqrt{\tau/N} \leq 2C(d,\omega)L\sqrt{\tau/N}.$$

Proof (cont.)

It follows by induction that

$$\|S^n[N]V - S^n V\| \leq C(d,\omega) nL\sqrt{\tau/N} = C(d,\omega) tL\sqrt{1/(\tau N)}, \tag{104}$$

yielding (102).

(ii) One shows as above that for any Lipschitz continuous policy π, the corresponding value functions $V^{\pi,N}$ converge. Combined with (i), this yields Statement (ii).

Infinite horizon: discounted payoff

The corresponding infinite horizon planning is the problem of maximizing the discounted sum

$$\Pi^{\pi,N}(x(N)) = \mathbf{E}_{N,x(N)} \sum_{k=0}^{\infty} \tau \beta^k B(x_k, b_k), \tag{105}$$

with a $\beta \in (0,1)$, where, as above,

$$x_k = X_N(\tau, x_{k-1}, b_{k-1}), \quad k = 1, 2, \cdots ,$$

$\pi = \{b_k\}$ and $b_k = b_k(x_k)$ are measurable functions depending on the current state $x = x_k$.
In the LLN limit, $N \to \infty$, we expect the limiting problem for the principal to be the maximization of the reward

$$\Pi^{\pi}(x) = \sum_{k=0}^{\infty} \beta^k \tau B(x_k, b_k) \tag{106}$$

with $x_k = X(\tau, x_{k-1}, b_{k-1})$.

Infinite horizon: discounted payoff

Notice firstly that the solution to the finite-time discounting problem of the maximization of the payoff

$$V_n^{\pi,N}(x(N)) = \mathbf{E}_{N,x(N)}\big[\tau B(x_0, b_0)$$

$$+ \cdots + \beta^{n-1}\tau B(x_{n-1}, b_{n-1}) + \beta^n V(x_n)\big], \tag{107}$$

is given by the iterations

$$V_n^N = S_\beta[N] V_{n-1}^N = S_\beta^n[N] V \tag{108}$$

of the corresponding discounted Shapley operator

$$S_\beta[N]V(x) = \sup_b \left[\tau B(x, b) + \beta \mathbf{E} V(X_N(\tau, x, b))\right]. \tag{109}$$

Infinite horizon: discounted payoff

Similarly the solution to the corresponding limiting discounted problem

$$V_n(x) = \max_{\pi} \left[\tau B(x_0, b_0) + \cdots + \beta^{n-1} \tau B(x_{n-1}, b_{n-1}) + \beta^n V(x_n) \right], \tag{110}$$

is given by the iterations

$$V_n = S_\beta V_{n-1}^N = S_\beta^n V \tag{111}$$

of the corresponding discounted Shapley operator

$$S_\beta V(x) = \sup_b \left[\tau B(x, b) + \beta V(X(\tau, x, b)) \right]. \tag{112}$$

Infinite horizon: discounted payoff

Recall a standard fact about the finite-step approximations to the optimal Π^π.

Proposition. (On finite-step approximations). Assume

$$\omega = \sup_b \|Q(., b)\|_{bLip} < \infty.$$

Let B and V be bounded continuous functions. Then the sequence $S_\beta^n V(x)$ converges, as $n \to \infty$, to the discounted infinite horizon optimal reward

$$\Pi(x) = \sup_\pi \Pi^\pi(x),$$

and the sequence $S_\beta^n[N]V(x(N))$ converges, as $n \to \infty$, to the discounted infinite horizon optimal reward

$$\Pi^N(x(N)) = \sup_\pi \Pi^{\pi,N}(x(N)).$$

Proof

Since

$$\|S_\beta^n V - S_\beta^n \tilde{V}\| \le \beta^n \|V - \tilde{V}\|,$$

$$\|S_\beta^n[N] V - S_\beta^n[N] \tilde{V}\| \le \beta^n \|V - \tilde{V}\|,$$

it follows that, if the sequences $S_\beta^n V$ or $S_\beta^n[N]V$ converge for some V, then these iterations converge to the same limit for any bounded V. But for $V = 0$ we see directly that

$$\|S_\beta^n V - \Pi\| \le \beta^n \|B\| \frac{2}{1-\beta},$$

$$\|S_\beta^n[N] V - \Pi^N\| \le \beta^n \|B\| \frac{2}{1-\beta}.$$

Theorem 2 (Convergence for infinite horizon discounted payoff)

Assume (100) and (95) hold and let

$$\beta e^{\tau\omega} \leq \beta_0 < 1.$$

Then the discounted optimal rewards

$$\Pi^N(x(N)) = \sup_\pi \Pi^{\pi,N}(x(N))$$

converge, as $N \to \infty$ and $x(N) \to x$, to the discounted best reward

$$\Pi(x) = \sup_\pi \Pi^\pi(x).$$

Proof

Similarly to above,

$$\|S_\beta^n V\| \le \tau \|B\| \frac{1}{1-\beta_0} + \beta_0^n \|V\|,$$

$$\|S_\beta^n V\|_{Lip} \le \tau \varkappa_B (1 + \beta e^{\tau\omega} + \cdots$$
$$+ \beta^{n-1} e^{\tau\omega(n-1)}) + \beta^n \varkappa_V e^{\tau\omega n}$$
$$\le \tau \varkappa_B \frac{1}{1-\beta_0} + \beta_0^n \varkappa_V,$$

that is, unlike optimization without discounting, these norms are bound uniformly in the number of steps used.

Proof (cont.)

Estimating the differences $\|S_\beta^n[N]V - S_\beta^n V\|$ as in the proof of the previous Theorem yields

$$\|S_\beta^n[N]V - S_\beta^n V\|$$

$$\leq C\frac{\sqrt{\tau}}{\sqrt{N}}(\beta\|S_\beta^n V\|_{bLip} + \cdots + \beta^n\|S_\beta V\|_{bLip})$$

$$\leq C\frac{\sqrt{\tau}}{\sqrt{N}}\left(\tau\|B\|_{bLip}\frac{1}{(1-\beta_0)^2} + \beta_0^{n+1}\|V\|_{bLip}\right).$$

Since $S_\beta^n V(x(N))$ converges to $\Pi(x)$, as $n \to \infty$, it follows that $S_\beta^n[N]V(x)$ converges to $\Pi(x)$, as $n \to \infty$ and $N \to \infty$.

Remark. The optimal payoffs $\Pi(x)$ and $\Pi^N(x)$ are the fixed points of the Shapley operator: $S_\beta \Pi = \Pi$, $S_\beta[N]\Pi^N = \Pi$. This fact can be used as a basis for another proof of the Theorem.

Continuous time modeling

Let the state space of the group be again given by vectors $x = (n_1, \cdots, n_d)/N$ from the lattice $\mathbf{Z}_+^d/N$, but the efforts (budget) b of the major player be chosen continuously in time aiming at optimizing the payoff

$$\int_t^T B(x(s), b(s))\, ds + S_T(x(T))$$

where B, S_T are some continuous functions, uniformly Lipschitz in all their variables. The optimal payoff of the major player is thus

$$S_N(t, x(N)) = \sup_{b(.)} \mathbf{E}_{x(N)}^N \left\{ \int_t^T B(x(s), b(s))\, ds + S_T(x(T)) \right\}, \tag{113}$$

where $\mathbf{E}_x^N$ is the expectation of the corresponding Markov process starting at the position (x) at time t, and $\tilde{U}$ is some class of controls (say, piecewise constant).

Continuous time modeling

We are in the standard setting of a controlled Markov process generated by the operator $L_{b,N}$ from (90).
As was shown above, the operators $L_{b,N}$ tend to a simple first order PDO, so that the limiting optimization problem of the major player turns out to be the problem of finding

$$S(t,x) = \sup_{b(.)} \left\{ \int_t^T B(x(s), b(s))ds + S_T(x(T)) \right\}, \quad (114)$$

where $x(s)$ solve the system of equations

$$\dot{x}_j = \sum_i x_i Q_{ij}(x,b), \quad j = 1, ..., d. \quad (115)$$

Continuous time modeling

The optimal payoff $S(t,x)$ of (114) represents the unique generalized solution to the HJB equation

$$\frac{\partial S}{\partial t} + \sup_b \left[B(x,b) + \left(\frac{\partial S}{\partial x}, xQ(x,b) \right) \right] = 0, \qquad (116)$$

with the initial (or terminal) condition $S(T,x) = S_T(x)$.

Instead of proving the convergence $S_N(t, x(N)) \to S(t,x)$, we shall concentrate on a more practical issue comparing the corresponding discrete time approximations, as these approximations are usually exploited for practical calculations of S_N or S.

Continuous time modeling

The discrete-time approximation to the limiting problem of finding (114) is the problem of finding

$$V_{t,n}(x) = \sup_{\pi} V^{\pi}_{t,n}(x)$$

$$= \sup_{\pi} \left[\tau B(x_0, b_0) + \cdots + \tau B(x_{n-1}, b_{n-1}) + V(x_n)\right], \quad (117)$$

where $\tau = (T - t)/n$, $x_0 = x$, $V(x) = S_T(x)$ and

$$x_k = X(\tau, x_{k-1}, b_{k-1}), \quad k = 1, 2, \cdots, \quad (118)$$

with $X(t, x, b)$ solving equation (97) with the initial condition x at time $t = 0$.

Continuous time modeling

The discrete-time approximation to the initial optimization problem is the problem of finding

$$V_{t,n}^{N}(x_0) = \sup_{\pi} V_{t,n}^{\pi,N}(x_0)$$

$$= \sup_{\pi} \mathbf{E}_{N,x(N)} \left[\tau B(x_0, b_0) + \cdots + \tau B(x_{n-1}, b_{n-1}) + V_0(x_n) \right], \tag{119}$$

where $x_k = X_N(\tau, x_{k-1}, b_{k-1})$ with $X_N(t, x, b)$ denoting the Markov process with generator (90) and with the strategies $\pi = \{b_k\}$.

Convergence to HJB (remarks)

It is known (see e. g. Theorem 3.4 of V. N. Kolokoltsov and V. P. Maslov *Idempotent analysis and its applications*, Kluwer Publishing House, 1997) that the discrete approximations $V_n(x)$ approach the optimal solution $S(T-t,x)$ given by (114) and solving the Cauchy problem for the HJB (116).

It is also known (see e. g. Theorem 4.1 of W. H. Fleming, H. M. Soner *Controlled Markov Processes and Viscosity Solutions*, Sec. Ed. Springer 2006), that $V_n^N(x)$ with $V_0 = S_T$ approach the optimal solutions $S_N(T-t,x)$ given by (113) and solving certain HJB equation.

Theorem 3 (on Continuous time modeling)

Let B, S_T be uniformly Lipschitz in all their variables. Then, for any x, $t \in [0, T]$ and $\tau = N^{-\epsilon}$ with $\epsilon \in (0, 1)$,

$$|V^N_{t,n}(x) - V_{t,n}(x)| \leq C(d, \omega, T)(\varkappa_B + \varkappa_V)(N^{-(1-\epsilon)/2} + |x(N) - x|). \tag{120}$$

And consequently $V^N_{t,n}(x)$ converge, as $N \to \infty$ (and $n = (T - t)/\tau$, $\tau = N^{-\epsilon}$), to the optimal solution $S(T - t, x)$ given by (114) and solving the Cauchy problem for the HJB (116).

Proof: a direct consequence of Theorem 1.

Remarks on further developments

Two major players zero-sum game,

Several major players,

Stationary solutions and the turnpike theory.

Chapter 4

Mean-field games (MFGs) for finite-state models.

MFG consistency problem, approximate Nash equilibrium for a finite number of players, forward-backward systems, master equation, MFGs with a major player.

The difference between the MFG setting and the modeling above is that now the small agents themselves become rational optimizers and are not supposed just to follow some prescribed deterministic or stochastic strategies (like myopic behavior).

For simplicity, we treat MFG exclusively for finite state models.

Controlled Markov chains

Let $Q = \{Q_{ij}(t, u)\}$ be a family of Q-matrices, $i, j \in \{1, \cdots, d\}$, depending continuously on $u \in U \subset \mathbf{R}^n$ and piecewise continuously on t.
Markov strategy: a piecewise continuous functions $\hat{u}_j(t)$.
A controlled Markov chain: $X^{\hat{u}}_{s,i}(t)$ with the Q-matrices $Q_{ij}(t, \hat{u}_i(t))$.

Suppose one gets a payoff $J(t, j, u)$ per unit of time staying at j with control u around time t and the terminal payoff $S_T(j)$, paid at time T if the process terminates in j at time T.
The *Markov control problem*: to find the maximal total payoff

$$S(t, j) = \max_{\hat{u}} \mathbf{E}[\int_t^T J(s, j(s), \hat{u}_{j(s)}(s))ds + S_T(X^{\hat{u}}_{t,x}(T))],$$

where $j(s) = X^{\hat{u}}_{t,j}(s)$, and a strategy, called an *optimal strategy*, where this maximum is attained.

Controlled Markov chains

The *Bellman equation*:

$$\frac{\partial S(t,j)}{\partial t} + \max_u [J(t,j,u) + \sum_{k=1}^{d} Q_{jk}(t,u) S(t,k)] = 0, \quad (121)$$

with the terminal condition $S(T,j) = S_T^j$.

Heuristic derivation: Assuming S is smooth, and taking into account that $Q_{ij}\tau$ is approx. the probability to jump from i to j in a small time τ, yields approximately (optimality principle)

$$S(t,j) = \max_u [J(t,j,u)\tau$$

$$+\tau |Q_{jj}| \sum_{k \neq j} \frac{Q_{jk}}{|Q_{jj}|} S(t+\tau, k) + (1 - \tau |Q_{jj}|) S(t+\tau, j)].$$

Expanding S in the Taylor series and keeping only terms linear in τ (free terms $S(t,j)$ cancel) yields (121).

Remark. Rigorous justification: *verification theorem*.

MFG setting

Suppose we have a family $Q = \{Q_{ij}(x, u)\}$ of Q-matrices, $i, j \in \{1, \cdots, d\}$, depending continuously on parameter $u \in U$ and Lipshitz continuously on $x \in \Sigma_d$. Suppose there are N players, each moving according to Q and aiming at maximizing the payoff

$$\mathbf{E}[\int_t^T J(s, j(s), x(s), u(s))\, ds + S_T(j(T))], \tag{122}$$

where $j(s)$ is the position of the player at time s. Here the motions of all players are coupled, as all transitions depend on the total distribution of players $x = (n_1, \cdots, n_d)/N$, where n_j denote the number of players in states j, at any given time.

MFG methodology, I

It suggests the following solution concept for this problem for large N. Let an evolution of the distributions $x(t)$, $t \in [0, T]$, is a known continuous curve. Then any given player should search for the maximal payoff

$$S(t,j) = \max_{\hat{u}} \mathbf{E}[\int_t^T J(s, j(s), x(s), \hat{u}_{j(s)}(s))\, ds + S_T(j(T))],$$

where $j(s) = X^{\hat{u}}_{s,j}(s)$ is the Markov chain with the Q-matrices $Q_{ij}(x(t), \hat{u}_i(t))$. As follows from (121), $S(t,i)$ should satisfy the backward Bellman equation

$$\frac{\partial S(t,j)}{\partial t} + \max_u [J(t,j,x(t),u) + \sum_k Q_{jk}(x(t),u) S(t,k)] = 0, \tag{123}$$

with the terminal condition $S(T,j) = S_T(j)$.

MFG methodology, II

Assume we can find a solution $S(t,x)$ and the corresponding optimal strategy $\hat{u}_j(t) = \hat{u}_j(t, x(t))$ providing max in this equation at any time t. Now, if all players are using this optimal strategy, the mean-field interacting particle system of N players given by the Q-matrices $\hat{Q}_{ij}(t,x) = Q_{ij}(x, \hat{u}_i(t))$ converge, according to the main theorem of Chapter 2, to the solutions $X_{0,x(0)}(t)$ of the system of (forward) kinetic equations (45):

$$\dot{x}_k = \sum_{i \neq k} [x_i Q_{ik}(x, \hat{u}_i(t))) - x_k Q_{ki}(x, \hat{u}_k(t)))] = \sum_{i=1}^{d} x_i Q_{ik}(x, \hat{u}_i(t))), \tag{124}$$

with $k = 1, ..., d$.

MFG methodology, III

Let $\hat{x}(t)$ be the solution of this system with the initial condition $\hat{x}(0) = x(0)$. The consistency between the controlled dynamics and the mean-field evolution can be naturally described by the requirement that $\hat{x}(t) = x(t)$. This is exactly the forward-backward *MFG consistency condition*, also referred by some authors as *Nash-MFG equilibrium*.

Equivalently starting with a control $u_j(t)$ we can solve the corresponding kinetic equation finding the distribution $x(t)$ and then find the corresponding optimal control $\hat{u}_j(t)$ fitting to the solution of the HJB (123). The MFG consistency condition can then be expressed by the equation $u(t) = \hat{u}(t)$.

MFG methodology, IV

In any case it can be expressed by saying that the pair $(\hat{x}(t), \hat{u}_j(t))$ provides a solution to the coupled *forward-backward system* (123)-(124), more precisely, to its *initial-terminal value problem* (initial x_0 is given for (124) and terminal S_T for (123)).

One can expect that solutions to MFG consistency problem should provide some approximations to the solutions of the games of N players, where each player is trying to maximize (122).

Main MFG problems

One can expect that solutions to MFG consistency problem should provide some approximations to the solutions of the games of N players, where each player is trying to maximize (122).
Thus we are led to two basic problems of MFG theory: describe the solutions to the MFG consistency problem (say, prove existence and/or uniqueness theorem) and provide exact link between these solutions and the corresponding original game of a finite number of players.

The latter task can be discussed with two approaches (often requiring different techniques): showing that Nash equilibria of games with N players converge to a solution of the MFG consistency problem, or showing that the solutions of the MFG consistency problem yield approximate Nash equilibria for finite-player games.

Master equation

An alternative approach to the MFG consistency problem arises from looking directly at the limiting evolution of the pair $(j(t), x(t))$, where $x(t)$ evolves according to the kinetic equations,

$$\dot{x}_k = \sum_{i=1}^{d} x_i Q_{ik}(x(t), \hat{u}_i(t)), \quad k = 1, ..., d,$$

and $j(t) = X_{s,x}^{\hat{u}}(t)$ is the Markov chain with the Q-matrices $Q_{ij}(x(t), \hat{u}_i(t))$, as at the controlled Markov process in the continuous state space $\{1, \cdots, d\} \times \Sigma_d$.

Master equation

This is, strictly speaking, not a chain anymore, as it evolves by jumps and continuous displacements in the continuous state space. Nevertheless, controlling the process with the objective to maximize (122) we find for the optimal payoff the Bellman equation in the form

$$\frac{\partial S(t,j,x)}{\partial t} + \max_{u_1,\cdots,u_d}\Big[J(t,j,x,u_j) + \sum_k Q_{jk}(x,u_j)S(t,k,x) + \sum_{k,i}\frac{\partial S(t,j,x)}{\partial x_k} x_i Q_{ik}(x,u_i)\Big] = 0. \qquad (125)$$

Master equation

This is obtained analogously to (121) from the approximate equation

$$S(t,j,x) = \max_{u_1,\cdots,u_d} [\tau J(t,j,x,u_j)$$

$$+\tau|Q_{jj}|\sum_{k\neq j}\frac{Q_{jk}(x,u_j)}{|Q_{jj}|}S(t+\tau,k,X_{x,t}(t+\tau))$$

$$+(1-\tau|Q_{jj}|)S(t+\tau,j,X_{x,t}(t+\tau))].$$

Master equation

Discarding the higher terms in τ,

$$S(t,j,x) = \max_{u_1,\cdots,u_d}[J(t,j,x,u)\tau + \tau\sum_{k\neq j} Q_{jk}S(t,k,x)$$

$$+(1-\tau|Q_{jj}|)S(t+\tau,j,X_{x,t}(t+\tau))].$$

Using the first order Taylor approximation

$$S(t+\tau,j,X_{x,t}(t+\tau)) = S(t,j,x) + \tau\frac{\partial S}{\partial t}(t,j,x)$$

$$+\tau\sum_{k,i}\frac{\partial S}{\partial x_k}(t,j,x)x_i Q_{ik}(x,u_i).$$

yields (125).

Master equation

Equation (125) is called the *master equation* (in backward form). The next statement shows that this equation provides (at least in case it is reasonably well-posed) an alternative approach to the analysis of the MFG consistency problem, which selects the most effective solutions to backward-forward systems, thus forming an envelop for various solutions of the MFG consistency problem.

Theorem 1.

Let $S(t,j,x)$ be a smooth solution to (125) with the terminal condition $S_T(j)$ (giving the optimal payoff in the corresponding Markov decision problem on $\{1,\cdots,d\}\times\Sigma_d$). Let it be possible to choose Lipschitz continuous (in x) selectors $\tilde{u}_j(t,x)$ giving maximum in (125) (for this S) and hence to build the trajectories $X_{x_0}(t)$ solving the kinetic equations

$$\dot{x}_j = \sum_i x_i Q_{ij}(x, \tilde{u}_i(t,x))$$

with any initial x_0.

Then the pair $(X_{x_0}(t), \tilde{u}_j(t, X_{x_0}(t))$ is a solution to the MFG consistency problem and

$$\tilde{S}(t,j) \leq S(t,j,x_0), \tag{126}$$

for payoff $\tilde{S}(t,j)$ on any solution $(\hat{x}(t), \hat{u}_j(t))$ to the forward-backward MFG consistency problem with $\hat{x}(t) = X_{x_0}(t)$.

Proof

By the definition of $S(t, j, x)$ as the solution to a Markov decision problem in $\{1, \cdots, d\} \times \Sigma_d$, $S(t, j, x_0)$ is not less than the payoff that can be obtained by any player using any symmetric strategy (the same as all other players) given the dynamics $X_{x_0}(t)$ of the total distributions. Thus (126) follows.

It follows that the payoff $S(t, j, x_0)$ cannot be improved by changing the strategies inside the class of symmetric strategies. Consequently, $\tilde{u}_i(t, x(t))$ provides the maximum payoff in the class of these strategies and hence provides a solution to the forward-backward MFG consistency problem.

MFG solutions and Nash equilibria for finite number of player

The Nash -equilibria and the ϵ-Nash equilibria for dynamic N-player game can be looked at in several ways, which are traditionally distinguished in the literature on optimization theory.

In general, one speaks about *open loop control* and related *open loop equilibria*, if players choose their control strategies $u_j(t)$ from the beginning, irrespectively of the dynamics of the game (but which may depend on the common source of uncertainty).

One speaks about *closed loop control* and related *closed loop equilibria*, if players choose *feedback controls* $u_j(t, z)$, which at any time t depends also on the position z of the process.

MFG solutions and Nash equilibria for finite number of player

In MFG setting new possibilities arise, as from the point of view of each player the position incudes his/her own position, say j, and the overall distribution x.
Let us speak about *partially open loop control* and related *partially open loop equilibria* if each player chooses among strategies $u_j(t)$, piecewise continuous in t, that depend on his/her own position at time t, but not on the overall distribution x. The use of such strategies, sometimes referred to as *distributed strategies*, is reasonable in many cases, where the overall distribution is not easily observable by each concrete player.
By a *close loop control* we mean, as usual, control $u_j(t, x)$ that is supposed to be applied by a player at time t when his/her position is j and the overall distribution is x.

MFG solutions and Nash equilibria for finite number of player

Let us concentrate on the case with the running cost function J not depending on x explicitly, that is, with the payoff

$$\mathbf{E}[\int_t^T J(s, j(s), u_{j(s)}(s))\, ds + S_T(j(T))], \tag{127}$$

with a continuous function J, and with the partially open loop equilibria. Thus we assume that we have a family $Q = \{Q_{ij}(x, u)\}$ of Q-matrices, $i, j \in \{1, \cdots, d\}$, depending continuously on parameter $u \in U$ and Lipschitz continuously on $x \in \Sigma_d$. Suppose there are N players, each moving according to Q and aiming at maximizing the payoff (127).

MFG solutions and Nash equilibria for finite number of player

Theorem 2.
Let $(\hat{x}(t), \hat{u}_j(t))$ be a solution to the backward-forward MFG consistency problem. Then, for the initial distribution $x(0)$, the symmetric profile of strategies $\hat{u}_j(t))$ is an ϵ-Nash equilibrium in the partially open loop setting, with ϵ of order $1/\sqrt{N}$. If $Q(x, u)$ is twice continuously differentiable in x uniformly in u, then the order of ϵ improves to $1/N$.

MFG solutions and Nash equilibria for finite number of player

Proof. We have to show that if all players use the strategy $\hat{u}_j(t)$, then any particular player unilaterally deviating from this strategy cannot increase the payoff by amount exceeding ϵ. Thus let us assume that one tagged player is using some deviating strategy $u_j^{dev}(t)$, while other players stick to $\hat{u}_j(t)$. The natural state space for such Markov chain will be $\{1, \cdots, d\} \times \Sigma_d$, the first discrete coordinate j denoting the position of the tagged player.

MFG solutions and Nash equilibria for finite number of player

We are exactly in the setting of Chapter 2. The operator $L_t^{N,dev}$ given by (75) and the limiting operator (76) take the form

$$L_t^{N,dev} f(j,x) = \sum_k Q_{jk}^{dev}(x, u_j^{dev}(t))(f(k,x) - f(j,x))$$

$$+\sum_i (x_i - \delta_i^j/N) \sum_{k \neq i} Q_{ik}(x, \hat{u}_j(t)) \left[f(j, x - e_i/N + e_k/N) - f(j,x) \right], \tag{128}$$

$$\Lambda_t^{dev} f(j,x) = \sum_k Q_{jk}^{dev}(x, u_j^{dev}(t))(f(k,x) - f(j,x))$$

$$+ \sum_i x_i \sum_{k \neq i} Q_{ik}(x, \hat{u}_j(t)) \left[\frac{\partial f}{\partial x_k} - \frac{\partial f}{\partial x_i} \right] (j,x). \tag{129}$$

MFG solutions and Nash equilibria for finite number of player

Applying Theorem 5 of Chapter 2 we derive that the payoffs for the tagged player in the N player game differs by his/her payoff in the limiting evolution by amount not exceeding ϵ. Notice that, in order to take into account the running payoff J, we apply this theorem not only to $S_T(j)$ and terminal time T, but also to each $J(s, j(s), u_{j(s)}(s))$ with terminal time s. Since $\hat{u}_j(t)$ is optimal in the limiting game it is therefore ϵ-optimal for the N player game.

Remark. To extend the theorem to the closed loop control and J depending on x, one just has to apply Theorem 6 that extends Theorem 5 to functions f depending explicitly on x.

MFG with a major player

Extension with a major player.
Let us assume that the transition matrices $Q = Q(x, u, b)$ and the payoffs $J(t, j, x, u, b)$ depend additionally on a parameter b controlled by the major player (a principal).
If the principal is playing just the best response $b^*(x)$ we are directly back to the original problem with $Q = Q(x, u, b^*(x)$.
However, if the major player choose b strategically aiming at maximizing some payoff of the general type

$$\int_t^T B(x(s), b(s))\, ds + V_T(x(T)),$$

the situation becomes different.

MFG with a major player

If the evolution of the distributions $x(t)$, $t \in [0, T]$, is a known continuous curve, the major player finds the optimal strategy $\hat{b}(t)$ and, based on this strategy, any given player should search for the maximal payoff

$$S(t, i) = \max_{\hat{u}} \mathbf{E}[\int_t^T J(s, j(s), x(s), \hat{u}_{j(s)}(s), \hat{b}(s))dt + S_T(j(T))],$$

where $j(t) = X_{s,x}^{\hat{u}}(t)$ is the Markov chain with the Q-matrices $Q_{ij}(x(t), \hat{u}_i(t), \hat{b}(t))$. The Bellman equation for the optimal payoff of each small player $S(t, i)$ takes the form

$$\frac{\partial S(t, j)}{\partial t} + \max_u [J(t, j, x(t), u, \hat{b}(t))$$

$$+ \sum_k Q_{jk}(x(t), u, b(t))S(t, k)] = 0.$$

MFG with a major player

After finding a solution $S(t,x)$ and the corresponding optimal strategy $\hat{u}_j(t) = \hat{u}_j(t, x(t))$ providing max in this equation at any time t we can solve the corresponding kinetic equations

$$\dot{x}_k = \sum_{i=1}^{d} x_i Q_{ik}(x, \hat{u}_i(t)), \hat{b}(t)), \quad k = 1, ..., d.$$

Let $\hat{x}(t)$ be the solution of this system with the initial condition $\hat{x}(0) = x(0)$. The *MFG consistency problem* or *MFG consistency condition with major player* can be expressed by the equation $\hat{x}(t) = x(t)$.

Chapter 5

The simplest (fully solvable) 3-state model of MFGs.

Inspection, corruption, botnet defence.

Three state model of corruption and inspection

A simple model of corruption that takes into account the effect of the interaction of a large number of agents by both rational decision making and myopic behavior is developed.

It describes the distribution of corrupted and honest agents under the pressure of both an incorruptible governmental representative ('benevolent principal') and the 'social norms' of the society.

Its stationary version turns out to be a rare example of an exactly solvable model of mean-field-game type. In particular, it reveals explicitly the non-uniqueness of solutions.

The results show clearly how the presence of interaction (including social norms) influences the spread of corruption by creating phase transition from one to three equilibria.

The model

Suppose an agent can be in one of the three states: honest H, corrupted C, reserved R, where R is the reserved job of low salary that an agent receives as a punishment if her corrupted behavior is discovered.

The change between H and C is subject to the decisions of the agents (though the precise time of the execution of their intent is noisy) the change from C to R are random with distributions depending on the level of the efforts (say, a budget used) b of the principal (a government representative) invested in chasing a corrupted behavior.

The change R to H (so-to-say, a new recruitment) may be possible and is included as a random event with a certain rate.

The model

Let n_H, n_C, n_R denote the numbers of agents in the corresponding states, $N = n_H + n_C + n_R$ the total number. States are 3-vectors $n = (n_H, n_C, n_R)$ or normalized versions $x = (x_H, x_C, x_R) = n/N$.

The control parameter u of each player in states H or C may have two values, 0 and 1, meaning that the player is happy with her state (H or C) or she prefers to switch one to another.

There is no control in the state R. When the updating decision 1 is made, the updating occurs with a rate λ. The recovery rate (rate of change from R to H) (recruited the agents start by being honest) is a constant r.

The model

Apart from taking a rational decision to swap H and C, an honest agent can be pushed to become corruptive by her corruptive peers, the effect being proportional to the fraction of corrupted agents with certain coefficient q_{inf}, which is analogous to the infection rate in epidemiologic models.

On the other hand, the honest agents can contribute to chasing and punishing corrupted behavior, this effect of a desirable social norm being proportional to the fraction of honest agents with certain coefficient q_{soc}.

The model

Thus if all agents use the strategy $u_H, u_C \in \{0,1\}$ and the efforts of the principle is b, the evolution of the state x is clearly given by the ODE

$$\begin{cases} \dot{x}_R = (b + q_{soc} x_H) x_C - r x_R, \\ \dot{x}_H = r x_R - \lambda(x_H u_H - x_C u_C) - q_{inf} x_H x_C, \\ \dot{x}_C = -(b + q_{soc} x_H) x_C + \lambda(x_H u_H - x_C u_C) + q_{inf} x_H x_C. \end{cases} \tag{130}$$

Here u_H, u_C can be considered as arbitrary measurable functions of t.

The model

This system is a particular representative of kinetic equations (124). If all agents use the strategy $u_H, u_C \in \{0, 1\}$ and the efforts of the principal is b, the generator of the Markov evolution on the states n is

$$L_N F(n_H, n_C, n_R)$$

$$= n_C(b + q_{soc}\frac{n_H}{N})(F(n_H, n_C - 1, n_R + 1) - F(n_H, n_C, n_R))$$
$$+ n_R r(F(n_H + 1, n_C, n_R - 1) = F(n_H, n_C, n_R))$$
$$+ \lambda n_C u_C(F(n_H + 1, n_C - 1, n_R) - F(n_H, n_C, n_R))$$
$$+ n_H(\lambda u_H + q_{inf}\frac{n_C}{N})(F(n_H - 1, n_C + 1, n_R) - F(n_H, n_C, n_R)).$$

The model

For any N, this generator describes a Markov chain on the finite state space $\{n = (n_H, n_C, n_R) : n_H + n_C + n_R = N\}$, where any agent, independently of others, can be recruited with rate r (if in state R) or change from C to H or vice versa if desired (with rate λ), and where the change of the state due to binary interactions are taken into account by the terms containing q_{soc} and q_{inf}.

The model

In terms of x the generator $L_N F$ takes the form

$$L_N F(x) = x_C(b + q_{soc}x_H)(F(x - e_C/N + e_R/N) - F(x))$$

$$+x_R r(F(x - e_R/N + e_H/N) - F(x))$$

$$+\lambda x_C u_C(F(x - e_C/N + e_H/N) - F(x))$$

$$+ x_H(\lambda u_H + q_{inf} x_C)(F(x - e_H/N + e_C/N) - F(x)), \quad (131)$$

where $\{e_j\}$ is the standard basis in $\mathbf{R}^3$.

The model

If F is a differentiable function, $L_N F$ converges to

$$LF(x) = x_C(b + q_{soc}x_H)\left(\frac{\partial F}{\partial x_R} - \frac{\partial F}{\partial x_C}\right) + x_R r\left(\frac{\partial F}{\partial x_H} - \frac{\partial F}{\partial x_R}\right)$$

$$+ x_H(\lambda u_H + q_{inf}x_C)\left(\frac{\partial F}{\partial x_C} - \frac{\partial F}{\partial x_H}\right) + \lambda x_C u_C\left(\frac{\partial F}{\partial x_H} - \frac{\partial F}{\partial x_C}\right), \tag{132}$$

as $N \to \infty$, which follows from the Taylor formula. This is a first order partial differential operator and its characteristics are given by ODE (130).

The model

If $x(t)$ and $b(t)$ are given, the dynamics of each individual player is the Markov chain on the 3 states with the generator

$$\begin{cases} L^{ind}g(R) = r(g(H) - g(R)) \\ L^{ind}g(H) = (\lambda u_H^{ind} + q_{inf}x_C)(g(C) - g(H)) \\ L^{ind}g(C) = \lambda u_C^{ind}(g(H) - g(C)) + (b + q_{soc}x_H)(g(R) - g(C)) \end{cases} \tag{133}$$

depending on the individual control $u^{ind} \in \{0, 1\}$, so that $\dot{g} = L^{ind}g$ is the Kolmogorov backward equation of this chain.

The model

Assume that an employed agent receives a wage w_H per unit of time and, if corrupted, an average payoff w_C (that includes w_H plus some additional illegal reward); she has to pay a fine f when her illegal behavior is discovered; the reserved wage for fired agents is w_R.

Thus the total payoff for a player on the time period $[t, T]$ is $\int_t^T w_S(\tau)d\tau + fM(t, T)$, where S denotes the state (which is either R, or H, or C) and $M(t, T)$ is the number of transitions from C to R during the period.

The model

If the distribution of other players is $x(t) = (x_R, x_H, x_C)(t)$, the HJB equation of type (125) describing the expectation of the optimal payoff $g = g_t$ (starting at time t with time horizon T) of an agent is

$$\begin{cases} \dot{g}(R) + w_R + r(g(H) - g(R)) = 0 \\ \dot{g}(H) + w_H + \max_u (\lambda u + q_{inf} x_C)(g(C) - g(H)) = 0 \\ \dot{g}(C) + w_C - (b + q_{soc} x_H) f + \max_u (\lambda u (g(H) - g(C)) \\ \qquad + (b + q_{soc} x_H)(g(R) - g(C)) = 0. \end{cases} \tag{134}$$

Exercise. Using the methods of the end of Chapter 4 of Part 2 derive this system.

MFG consistency

Starting with a control

$$u^{com}(t) = (u_C^{com}(t), u_H^{com}(t)),$$

used by all players, we can find the dynamics $x(t)$ from equation (130) (with u^{com} used for u). Then each individual should solve the Markov control problem (134) thus finding the individually optimal strategy

$$u^{ind}(t) = (u_C^{ind}(t), u_H^{ind}(t)).$$

The *MFG consistency condition* can now be written as

$$u^{ind}(t) = u^{com}(t). \tag{135}$$

Stationary MFG consistency problem

Two standard stationary problems arising from HJB (134): for average payoffs and for discounted payoff.
Search for average payoff

$$g = \lim_{T\to\infty} \frac{1}{T} \int_0^T g_t \, dt$$

for long period games is governed by the solutions of HJB of the form $(T-t)\mu + g$, linear in t (with μ describing the optimal average payoff), so that g satisfies the stationary HJB:

$$\begin{cases} w_R + r(g(H) - g(R)) = \mu \\ w_H + \max_u (\lambda u + q_{inf} x_C)(g(C) - g(H)) = \mu \\ w_C - (b + q_{soc} x_H) f + \max_u (\lambda u (g(H) - g(C)) \\ \qquad + (b + q_{soc} x_H)(g(R) - g(C)) = \mu, \end{cases} \tag{136}$$

Stationary MFG consistency problem

The discounted optimal payoff is given by

$$g_t(S) = \sup\left(\int_t^\infty e^{-\delta\tau} w_S(\tau)d\tau + f\sum_j e^{-\delta t_j}\right),$$

where the sum is over all moment t_j, when a transition from C to R occurs. If g_t does no depend on time (stationary solution), then it satisfies the stationary HJB.

$$\begin{cases} w_R + r(g(H) - g(R)) = \delta g(R) \\ w_H + \max_u(\lambda u + q_{inf} x_C)(g(C) - g(H)) = \delta g(H) \\ w_C - (b + q_{soc} x_H) f + \max_u(\lambda u(g(H) - g(C)) \\ \qquad + (b + q_{soc} x_H)(g(R) - g(C)) = \delta g(C). \end{cases} \tag{137}$$

Exercise. Using the methods of the end of Chapter 4 of Part 2 derive the equation for g_t and show that a stationary solution satisfies (137).

Stationary MFG consistency problem

For this model we shall concentrate on the first problem (average payoff).

For a fixed b, the *stationary MFG consistency problem* for the average payoff is in finding $(x, u_C, u_H) = (x, u_C(x), u_H(x))$, where x is the stationary point of evolution (130), that is

$$\begin{cases} (b + q_{soc}x_H)x_C - rx_R = 0 \\ rx_R - \lambda(x_H u_H(x) - x_C u_C(x)) - q_{inf}x_H x_C = 0 \\ -(b + q_{soc}x_H)x_C + \lambda(x_H u_H(x) - x_C u_C(x)) + q_{inf}x_H x_C = 0, \end{cases} \tag{138}$$

where $u_C(x), u_H(x)$ are the maximizers in (136). Thus x is a fixed point of the limiting dynamics of the distribution of large number of agents such that the corresponding stationary control is individually optimal subject to this distribution.

Stationary MFG consistency problem

Fixed points can practically model a stationary behavior only if they are stable. Thus we are interested in stable solutions $(x, u_C, u_H) = (x, u_C(x), u_H(x))$ to the stationary MFG consistency problem, where a solution is stable if the corresponding stationary distribution $x = (x_R, x_H, x_C)$ is a stable equilibrium to (130) (with u_C, u_H fixed by this solution).

Major concern: to find out how the presence of interaction (specified by the coefficients q_{soc}, q_{inf}) affects the stable equilibria.

Solutions to the stationary problem

We shall say that in a solution to the stationary MFG consistency problem

the optimal individual behavior is corruption if $u_C = 0, u_H = 1$: if you are corrupt stay corrupt, and if you are honest, start corrupted behavior as soon as possible;

the optimal individual behavior is honesty if $u_C = 1, u_H = 0$: if you are honest stay honest, if you are involved in corruption try to clean yourself from corruption as soon as possible.

Solutions to the stationary problem

The natural assumptions on our coefficients, arising from the interpretation above, are:

$$\lambda > 0, r > 0, b > 0, \quad f \geq 0, q_{soc} \geq 0, q_{inf} \geq 0,$$

$$w_C > w_H > w_R \geq 0. \tag{139}$$

The key parameter for our model turns out to be the quantity

$$\bar{x} = \frac{1}{q_{soc}} \left[\frac{r(w_C - w_H)}{w_H - w_R + rf} - b \right] \tag{140}$$

(which can take values $\pm\infty$ if $q_{soc} = 0$).

Theorem

Assume (139).
(i) If $\bar{x} > 1$, then there exists a unique solution $x^* = (x_R^*, x_C^*, x_H^*)$ to the stationary MFG problem (138), (136), where

$$x_C^* = \frac{(1 - x_H^*)r}{r + b + q_{soc}x_H^*} \tag{141}$$

and x_H^* is the unique solution on the interval $(0, 1)$ of the quadratic equation $Q(x_H) = 0$, where

$$Q(x_H) = [(r+\lambda)q_{soc} - rq_{inf}]x_H^2 + [r(q_{inf} - q_{soc}) + \lambda r + \lambda b + rb]x_H - rb. \tag{142}$$

Under this solution the optimal individual behavior is corruption: $u_C = 0, u_H = 1$.

Theorem (cont.)

(ii) If $\bar{x} < 1$, there may be 1,2 or 3 solutions to the stationary MFG problem (138), (136). Namely, the point $x_H = 1, x_C = x_R = 0$ is always a solution, under which the optimal individual behavior is being honest: $u_C = 1, u_H = 0$.

Moreover, if

$$\max(\bar{x}, 0) \leq \frac{b + \lambda}{q_{inf} - q_{soc}} < 1, \tag{143}$$

then there is another solution with the optimal individual behavior being honest, that is $u_C = 1, u_H = 0$:

$$x_H^{**} = \frac{b + \lambda}{q_{inf} - q_{soc}}, \quad x_C^{**} = \frac{r(q_{inf} - q_{soc} - b - \lambda)}{(r + b)q_{inf} + (\lambda - r)q_{soc}}. \tag{144}$$

Theorem (cont.)

Finally, if

$$\bar{x} > 0, \quad Q(\bar{x}) \geq 0, \tag{145}$$

there is a solution with the corruptive optimal behavior of the same structure as in (i), that is, with x_H^* being the unique solution to $Q(x_H) = 0$ on $(0, \bar{x}]$ and x_C^* given by (141).

Remark. As seen by inspection, if $q_{inf} - q_{soc} > 0$, then $Q[(b+\lambda)/(q_{inf} - q_{soc})] > 0$, so that for $\bar{x}$ slightly less than $x_H^{**} = (b+\lambda)/(q_{inf} - q_{soc})$ one has also $Q(\bar{x}) > 0$, in which case one really has three points of equilibria given by $x_H^*, x_H^{**}, x_H = 1$ with $0 < x^* < \bar{x} < x^{**} < 1$.

Proof

Solutions to (136) are defined up to an additive constant. Thus we can and will assume that $g(R) = 0$. Moreover, we can reduce the analysis to the case $w_R = 0$ by subtracting it from all equations of (136) and thus shifting by w_R the values w_H, w_C, μ. Under these simplifications, the first equation in (136) is $\mu = rg(H)$, so that (136) becomes the system

$$\begin{cases} w_H + \lambda \max(g(C) - g(H), 0) + q_{inf} x_C (g(C) - g(H)) = rg(H) \\ w_C - (b + q_{soc} x_H) f + \lambda \max(g(H) - g(C), 0) \\ \qquad - (b + q_{soc} x_H) g(C) = rg(H) \end{cases} \tag{146}$$

for the pair $(g(H), g(C))$ with $\mu = rg(H)$.

Proof (cont.)

Assuming $g(C) \geq g(H)$, that is $u_C = 0, u_H = 1$, so that the corruptive behavior is optimal, system (146) turns to

$$\begin{cases} w_H + \lambda(g(C) - g(H)) + q_{inf} x_C (g(C) - g(H)) = rg(H) \\ w_C - (b + q_{soc} x_H) f - (b + q_{soc} x_H) g(C) = rg(H). \end{cases} \tag{147}$$

Proof (cont.)

Solving this system of two linear equations we get

$$g(C) = \frac{(r + \lambda + q_{inf} x_C)[w_C - (b + q_{soc} x_H) f] - r w_H}{r(\lambda + q_{inf} x_C + b + q_{soc} x_H) + (\lambda + q_{inf} x_C)(b + q_{soc} x_H)},$$

$$g(H) = \frac{(\lambda + q_{inf} x_C)[w_C - (b + q_{soc} x_H) f] + (b + q_{soc} x_H) w_H}{r(\lambda + q_{inf} x_C + b + q_{soc} x_H) + (\lambda + q_{inf} x_C)(b + q_{soc} x_H)}.$$

Exercise. Perform the calculations leading to this solution.

Proof (cont.)

Hence $g(C) \geq g(H)$ is equivalent to

$$w_C - (b + q_{soc}x_H)f \geq w_H \left(1 + \frac{b + q_{soc}x_H}{r}\right),$$

or, in other words,

$$x_H \leq \frac{1}{q_{soc}} \left[\frac{r(w_C - w_H)}{w_H + rf} - b\right], \tag{148}$$

which by restoring w_R (shifting w_C, w_H by w_R) gives

$$x_H \leq \bar{x} = \frac{1}{q_{soc}} \left[\frac{r(w_C - w_H)}{w_H - w_R + rf} - b\right]. \tag{149}$$

Since $x_H \in (0, 1)$, this is automatically satisfied if $\bar{x} > 1$, that is under the assumption of (i). On the other hand, it definitely cannot hold if $\bar{x} < 0$.

Theorem (cont.)

Assuming $g(C) \leq g(H)$, that is $u_C = 1, u_H = 0$, so that the honest behavior is optimal, system (146) turns to

$$\begin{cases} w_H + q_{inf} x_C (g(C) - g(H)) = rg(H) \\ w_C - (b + q_{soc} x_H) f + \lambda (g(H) - g(C)) \\ \qquad - (b + q_{soc} x_H) g(C) = rg(H). \end{cases}$$

Proof (cont.)

Solving this system of two linear equations we get

$$g(C) = \frac{(r + q_{inf}x_C)[w_C - (b + q_{soc}x_H)f] + (\lambda - r)w_H}{r(\lambda + q_{inf}x_C + b + q_{soc}x_H) + q_{inf}x_C(b + q_{soc}x_H)}$$

$$g(H) = \frac{q_{inf}x_C[w_C - (b + q_{soc}x_H)f] + (\lambda + b + q_{soc}x_H)w_H}{r(\lambda + q_{inf}x_C + b + q_{soc}x_H) + q_{inf}x_C(b + q_{soc}x_H)}$$

so that $g(C) \leq g(H)$ is equivalent to the inverse of condition (148).

Exercise. Perform the calculations leading to this solution.

Proof (cont.)

If $g(C) \geq g(H)$, that is $u_C = 0, u_H = 1$, the fixed point equation (138) becomes

$$\begin{cases} (b + q_{soc} x_H) x_C - r x_R = 0 \\ r x_R - \lambda x_H - q_{inf} x_H x_C = 0 \\ -(b + q_{soc} x_H) x_C + \lambda x_H + q_{inf} x_H x_C = 0. \end{cases} \tag{150}$$

Since $x_R = 1 - x_H - x_C$, the third equation is a consequence of the first two equations, which yields the system

$$\begin{aligned} &(b + q_{soc} x_H) x_C - r(1 - x_H - x_C) = 0 \\ &r(1 - x_H - x_C) - \lambda x_H - q_{inf} x_H x_C = 0. \end{aligned} \tag{151}$$

Proof (cont.)

From the first equation we have

$$x_C = \frac{(1 - x_H)r}{r + b + q_{soc}x_H}. \tag{152}$$

From this it is seen that if $x_H \in (0, 1)$ (as it should be), then also $x_C \in (0, 1)$ and

$$x_C + x_H = \frac{r + x_H(b + q_{soc}x_H)}{r + b + q_{soc}x_H} \in (0, 1).$$

Plugging x_C in the second equation of (151) we find for x_H the quadratic equation $Q(x_H) = 0$ with Q given by (142).

Proof (cont.)

Since $Q(0) < 0$ and $Q(1) > 0$, the equation $Q(x_H) = 0$ has exactly one positive root $x_H^* \in (0, 1)$. Hence x_H^* satisfies (148) if and only if either $\bar{x} > 1$ (that is we are under the assumption of (i)) or if (145) holds proving the last statement of (ii).

Proof (cont.)

If $g(C) \leq g(H)$, that is $u_C = 1, u_H = 0$, the fixed point equation (138) becomes

$$\begin{cases} (b + q_{soc} x_H) x_C - x_R r = 0 \\ x_R r + \lambda x_C - q_{inf} x_H x_C = 0 \\ -x_C (b + q_{soc} x_H) - \lambda x_C + q_{inf} x_H x_C = 0. \end{cases}$$

Proof (cont.)

Again here $x_R = 1 - x_H - x_C$ and the third equation is a consequence of the first two equations, which yields the system

$$\begin{cases} (b + q_{soc} x_H) x_C - r(1 - x_H - x_C) = 0 \\ r(1 - x_H - x_C) + \lambda x_C - q_{inf} x_H x_C = 0. \end{cases}$$

From the first equation we again get (152). Plugging this x_C in the second equation of (151) we find the equation

$$r(1 - x_H) = (r - \lambda + q_{inf} x_H) \frac{(1 - x_H) r}{r + b + q_{inf} x_H},$$

with two explicit solutions yielding the first and the second statements of (ii).

Further analysis

Stability of solutions in the sense of kinetic equations,

Stability in the sense of MFG (turnpikes),

The simplest 4-state model (botnet defence in cybersecuriy).

Lager state spaces, continuous state spaces, where our forward-backward ODEs turn to forward-backward PDEs.

Further applications: energy distribution, finance maths (order book modeling), Covid spreading, etc