Calculation of optima and equilibria in dynamic resource extraction problems (Research summary of doctoral dissertation)

Rajani Singh

Faculty of Mathematics, Informatics and Mechanics,
University of Warsaw
October 22, 2018

Game theory is a formal way of examining the situations of conflict and co-operation. A game is a mathematical tool to describe any situation in which there are at least two independent decision makers (called players), each of them has their own aim or objective (mathematically described as a maximization of a certain function called payoff), while there is a certain interdependence between them (mathematically described as dependence of the payoff function on choices of all the players). Formally it can be defined as

Definition 1. Game with finitely many players
A game in normal form for finitely many players $\mathcal{G} = \{\mathbb{I}, \{\mathbb{S}_i\}_{i\in\mathbb{I}}, \{J_i\}_{i\in\mathbb{I}}\}$ *consists of:*

- *A set of at least two players* $\mathbb{I}$*. For finitely many players* $\mathbb{I} = \{1, \dots, n\}$*.*
- *A set of strategies* $\mathbb{S}_i$ *that are available to player* i*. If* $s_i \in \mathbb{S}_i$ *denotes the strategy chosen by player* i*, then* $s = (s_1, s_2 \dots, s_n)$ *is called a strategy profile.*

  We denote the set of all strategy profiles *by* $\Sigma = \mathbb{S}_1 \times \mathbb{S}_2 \times \cdots \times \mathbb{S}_n$*.*
- *A set of payoff functions* $J = (J_1, J_2, \dots, J_n)$*, where* $J_i : \Sigma \to \mathbb{R}$ *is called the payoff function of player* i*.*

Notational convention:
For brevity of notation we will write $[s_i, s_{\sim i}]$ for a profile of strategies $s = (s_1, \dots, s_n)$, where $s_{\sim i}$ denotes the strategy of the remaining players. So, for a strategy $\sigma \in \mathbb{S}_i$ and a profile $\tilde{s} \in \Sigma$, the symbol $(\sigma, \tilde{s}_{\sim i})$ denotes the profile $\tilde{s}$ with i-th coordinate replaced by σ.

If the number of players in a real life game theoretic application is sufficiently large, they start behaving in such a way that is best described by the games with a continuum of players. More formally:

Definition 2. Game with continuum of players
A game in normal form $\mathcal{G} = \{\mathbb{I}, \mathcal{L}, \lambda, (\mathbb{S}, \mathfrak{S}), \{\mathbb{S}_i\}_{i\in\mathbb{I}}, \{J_i\}_{i\in\mathbb{I}}\}$ *for the continuum of players consists of:*

- *The* continuum of players *is the set of players* $\mathbb{I} = [0, 1]$ *with the Lebesgue measure* λ *on the* σ*-field of its Lebesgue measurable subsets* $\mathcal{L}$*. Thus, the space of players is the measure space* $(\mathbb{I}, \mathcal{L}, \lambda)$ *instead of only the set* $\mathbb{I}$*.*

- Sets of available strategies of player i, $\mathbb{S}_i$ *are all subsets of a certain set* $\mathbb{S}$ *on which* σ*- field of its measurable subsets* $\mathfrak{S}$*, its measurability is considered, denoted by* $\mathfrak{S}$*. We assume that* $\mathbb{S}_i \in \mathfrak{S}$*.*

  For a function $s : \mathbb{I} \to \mathbb{S}$ *with* $s_i \in \mathbb{S}_i$ *(for uniformity of notation, we write* s_i *instead of* $s(i)$*), we call* strategy profiles *only such measurable function.*

  As before, Σ *denotes the* set of all strategy profiles *but now obviously the definition of profile encompasses measurability.*

- *Payoff functions of player* i*,* $J_i : \Sigma \to \mathbb{R}$*. In majority of applications* J_i *are of specific form:*

  $J_i(s) = \mathcal{P}_i (s_i, u^s)$ *for a measurable function for some* $\mathcal{P}_i : \mathbb{S} \times \mathrm{Conv}\mathbb{S} \to \bar{\mathbb{R}}$ *and* $u^s = \int_{\mathbb{I}} s_j d\lambda(j)$*, usually called* the aggregate of s*, where* Conv $\mathbb{S}$ *denotes the* convex hull *of a set* $\mathbb{S}$.

The most important solution concept of the non-cooperative game theory is the Nash equilibrium.

Definition 3. *A strategy profile* $\bar{s}$ *is a* ***Nash equilibrium*** *for* n*-player game, iff for every player* $i \in \mathbb{I}$ *and for every strategy* $s_i \in \mathbb{S}_i$ *of player* i*,*

$$J_i ([s_i, \bar{s}_{\sim i}]) \leq J_i ([\bar{s}_i, \bar{s}_{\sim i}]) .$$

For continuum of player game, "every i" *is replaced by* "almost every i" *and* "some i" *by* "i in a set of non-zero measure"*.*

An important property of a strategy profile, which is rarely fulfilled by Nash equilibria but considered as one of the most important properties in the case when it is assumed that the players can make the decision together, is Pareto-optimality.

Definition 4. *A strategy profile* $\bar{s}$ *is* ***Pareto-optimal*** *for* n*-player game, if there is no profile such that*

$$J_i(s) \geq J_i(\bar{s}) \text{ for all } i \in \mathbb{I} \text{ and } J_i(s) > J_i(\bar{s}) \text{ for some } i.$$

For continuum of player game, "every i" *is replaced by* "almost every i" *and* "some i" *by* "i in a set of non-zero measure"*.*

In the dissertation, we are especially interested in a special Pareto-optimal profile called the *social optimum.*

Definition 5. *A strategy profile $\bar{s}$ is the* ***social optimum*** *:*
(a) in the n-players game iff

$$\bar{s} \in \operatorname*{Argmax}_{s\in\Sigma} \sum_{i=1}^{n} J_i(s).$$

(b) in the continuum of player game iff

$$\bar{s} \in \operatorname*{Argmax}_{s\in\Sigma} \int\limits_{[0,1]} J_i(s) \cdot d\lambda(i).$$

Dynamic games are the games of the particular structure with dependence on time and decision made in multiple time instants. They may be of a very complicated form, and they may be with complete or incomplete information. Dynamic games are the only appropriate tool to model decision-making problems by independent but coupled players in an external environment changing in response to their decisions.

Dynamic games for finitely many players
A *dynamic game* with n-players consists of the following:

- A set of *finitely many players* $\mathbb{I} = \{1, \dots, n\}$.
- A *time set* $\mathbb{T}$: either *discrete* $\mathbb{T} = \{0, 1, \dots, T\}$ for a finite time horizon T and $\mathbb{T} = \{0, 1, 2, \dots\}$ for the infinite time horizon or *continuous* $\mathbb{T} = [0, T]$ for a finite time horizon and $\mathbb{T} = [0, \infty)$ for the infinite time horizon. We denote the initial time by t_0.
- A set of possible *states of the system* (*state set* for short) $\mathbb{X} \subseteq \mathbb{R}^n$. A system is characterized at each time by a *state variable* $x \in \mathbb{X}$.
- A *potential trajectory* X of the state of the system is defined as $X : \mathbb{T} \cup \{T + 1\} \to \mathbb{X}$ for discrete time with finite time horizon T, $X : \mathbb{T} \to \mathbb{X}$ otherwise, with an *initial state* of the system $X(t_0) = x_0 \in \mathbb{X}$.
- *The equivalent of the control parameter* in dynamic game is called the *decision or action* of player $i \in \mathbb{I}$ at time t and is denoted by s_i.
- A set of decisions of player i is $\mathbb{D}_i \subseteq \mathbb{R}^{m_i}$ (with strategies being the set of functions $S_i : \mathbb{T} \times \mathbb{X} \to \mathbb{D}_i$, to be defined later).

  Preliminary *set of all decision profiles* is denoted by $\Delta = \mathbb{D}_1 \times \mathbb{D}_2 \times \cdots \times \mathbb{D}_n$.

- There is a *state dependent constraint* on decisions or actions of player i, given by the correspondence $D_i : \mathbb{X} \multimap \mathbb{D}_i$ with $D_i(x) \subseteq \mathbb{D}_i$, called the *correspondence of currently available decisions.*
- A decision profile $s \in \Delta$ available at state x, with $s_i \in D_i(x)$ is defined as $s = (s_1, \ldots, s_n)$.

  Information Structure

  Strategies that are available to players may have different information structure. Unlike in dynamic optimization problems, it is essential to be very precise about the information structure.

  We are interested in the form of strategies $S_i : \mathbb{T} \times \mathbb{X} \to \mathbb{D}_i$ that are measurable in the case of continuous time and fulfil one more condition, to be defined later. These are called in various papers *closed loop, closed loop no-memory, feedback or Markovian.*

  In some specific cases, $S_i : \mathbb{X} \to \mathbb{D}_i$, with the same ambiguous terminology.

  We use the later form of strategies only in the infinite time horizon case and when the functions and the correspondences stated in the problem are not directly dependent on time.

  Throughout the dissertation, we will use the term *feedback* (prevalent in most recent dynamic games literature).
- *Behaviour and evolution of the state variable*, given functions $S_i : \mathbb{T} \times \mathbb{X} \to \mathbb{D}_i$ and a strategy profile $S = (S_1, \ldots, S_n)$ is described by the following equation:

  a first order difference equation in discrete time

  $$X(t+1) = \phi\left(t, X(t), S(t, X(t))\right); X(t_0) = x_0, \tag{1}$$

  for the state transition function $\phi : \mathbb{T} \times \mathbb{X} \times \Delta \to \mathbb{X}$.

  a differential equation in continuous time

  $$\dot{X}(t) = \phi\left(t, X(t), S(t, X(t))\right); X(t_0) = x_0, \tag{2}$$

  for almost every t and for a state transition function $\phi : \mathbb{T} \times \mathbb{X} \times \Delta \to \mathbb{R}^n$.

  In continuous time, some regularity assumption is additionally needed for S (e.g., jointly measurable and Lipschitz in $\mathbb{X} \times \Delta$ for almost every t), guaranteeing that

  $$\forall\ (t_0, x_0) \in \mathbb{T} \times \mathbb{X}\ \exists \text{ a unique X which fulfils Eq. (2).} \tag{3}$$

  The unique trajectory which solves Eq. (1) or Eq. (2) for given $S : \mathbb{T} \times \mathbb{X} \to \Delta$ is called the *trajectory corresponding to S*. If we want to emphasize that

X is corresponding to S, we write X^S. If we also want to emphasize the dependency on the initial condition we write $X^S_{t_0,x_0}$ or $X^S_{x_0}$.

Generally, it cannot be a priori assumed that S is Lipschitz with respect to X, since discontinuous strategies may appear at Nash equilibria, so we just have the condition (4).

- In discrete time, the *set of profiles* of strategies is of the form $\Sigma = \mathbb{S}_1 \times \cdots \times \mathbb{S}_n$ ($\mathbb{S}_i$ being the set of functions $S_i : \mathbb{T} \times \mathbb{X} \to \mathbb{D}_i$, called the sets of strategies of player i) and it is a certain set of functions $S : \mathbb{T}\times\mathbb{X} \to \Delta$ which fulfil $S_i(t,x) \in D_i(x)$, while in continuous time it is a set of all measurable function $S : \mathbb{T} \times \mathbb{X} \to \Delta$ which fulfils $S_i(t,x) \in D_i(x)$ and

  $$\text{such that Eq. (2) has a unique absolutely continuous solution on } \mathbb{T}\cap[t_0,+\infty). \tag{4}$$

  If Σ fulfils $\Sigma = \mathbb{S}_1 \times \cdots \times \mathbb{S}_n$, then the set of trajectories corresponding to $S \in \Sigma$ is called the *set of admissible trajectories* and is denoted by $\mathfrak{X}$.

- *Instantaneous or current payoff* is a function $P : \mathbb{I}\times\mathbb{T}\times\mathbb{X}\times\Delta \to \mathbb{R}\cup\{-\infty\}$. We denote the function $P(i,\cdot,\cdot,\cdot)$ by P_i and it is called the *current or instantaneous payoff* of player i.

  For a finite time horizon T, we also consider the *terminal payoffs* $G_i^* : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$.

- We consider the *discounting* of the payoffs by a *discount factor* $\beta \in (0,1)$. For discrete time, $\beta = \frac{1}{1+r}$, while for continuous time, $\beta = e^{-r}$, for $r > 0$, called the *interest rate* in economics.

- A payoff function $J_i : \mathbb{T}\times\mathbb{X}\times\Sigma \to \mathbb{R} \cup \{-\infty\}$ of player i is equal to his/her instantaneous payoffs, discounted and summed over time.

  For a profile S, the payoff function in *discrete time* fulfils:

  $$J_i(t_0,x_0,S) = \sum_{t=t_0}^{T} \beta^{t-t_0} P_i\left(t, X(t), S(t,X(t))\right) + \beta^{T+1-t_0} G_i^*(X(T+1)) \tag{5a}$$

  for the finite time horizon T

  $$J_i(t_0,x_0,S) = \sum_{t=t_0}^{\infty} \beta^{t-t_0} P_i\left(t, X(t), S(t,X(t))\right) \tag{5b}$$

  for the infinite time horizon

  for X given by Eq. (1).

For a profile S, the payoff function in *continuous time* fulfils:

$$J_i(t_0, x_0, S) = \int_{t=t_0}^{T} \beta^{t-t_0} P_i\left(t, X(t), S(t, X(t))\right) dt + \beta^{T+1-t_0} G_i^*(X(T+1)) \tag{6a}$$

for the finite time horizon T

$$J_i(t_0, x_0, U) = \int_{t=t_0}^{\infty} \beta^{t-t_0} P_i\left(t, X(t), S(t, X(t))\right) dt \tag{6b}$$

for the infinite time horizon

for X given by Eq. (2).

We assume that the functions P_i, ϕ, G_i^* are measurable on $\mathbb{T} \times \mathbb{X} \times \Delta$ and $\phi(t, \cdot, \cdot)$ is Lipschitz continuous in $\mathbb{X} \times \mathbb{U}$.

We do not impose other direct constraints on the sets or the functions defined before, but we assume that $J_i(t_0, x_0, S)$ is *always well defined.*

Dynamic games for the continuum of players

Definition of dynamic games for the continuum of players are similar to Definition for n-players with the following changes:

- The *space of players* $(\mathbb{I}, \mathcal{L}, \lambda)$ for a set of players $\mathbb{I} = [0, 1]$ with a Lebesgue measure λ on the σ-field of its Lebesgue measurable subsets $\mathcal{L}$.
- The *set of decisions of player* i, $\mathbb{D}_i$ is $\mathcal{D}$ measurable subsets of a measurable space $(\mathbb{D}, \mathcal{D})$.
- *Currently available decisions* are $D_i(x)$ for $D_i : \mathbb{X} \multimap \mathbb{D}_i$ with $D_i(x) \in \mathcal{D}$.
- A *profile of decisions available* at state x is any measurable function $s : \mathbb{I} \to \mathbb{D}$ with $s_i \in D_i(x)$. For uniformity of notation, we write s_i instead of $s(i)$. The set of all profiles of decision is denoted by Δ.
- The *time set* is $\mathbb{R}_+$
- *Current payoffs* $P_i(t, X(t), s)$ are of specific form. They can be written as $\mathcal{P}_i(x, s_i, u^s)$, for some $\mathcal{P}_i : \mathbb{X} \times \mathbb{D} \times \text{Conv } \mathbb{D} \to \bar{\mathbb{R}}$, where $u^s = \int_{\mathbb{I}} s_j d\lambda(j)$, usually called *the aggregate of* s and Conv $\mathbb{D}$ denotes the convex hull of the set $\mathbb{D}$.
- The trajectory of the state variable corresponding to a profile of strategies S is $X^S(t+1) = \varphi(X^S(t), u^S(t))$ for a function $\varphi : \mathbb{X} \times \text{Conv } \mathbb{D} \to \mathbb{X}$ and $u^S(t) = u^{S(X(t))}$ with the initial condition $X(0) = x_0$.

- For a given profile S, the payoff function of player i is

$$J_i(x_0, S) = \sum_{t=0}^{\infty} \beta^t \mathcal{P}_i\big(X(t), S_i(X(t), u^S(t))\big).$$

The dynamic game which we mainly considered in the dissertation belongs to the class of linear-quadratic dynamic games with constraints. The real-life problems that are solved by using the tools of dynamic games and/or dynamic optimization in the dissertation are the model of extraction of a shared, renewable resource.

Extraction or exploitation of shared renewable resources is one of the most significant problems in society. It encompasses a wide range of various problems among other things, the phenomenon known as *the tragedy of the commons*. Most importantly, the extraction and consumption of common natural renewable resources have a strong impact on the quality of life and well-being of both, the current and future generations. From the mathematical point of view, the only tool to deal with the whole spectrum of phenomena arising in such types of problems, in which there are at least two independent decision makers in a common resource extraction problem, are dynamic games, since both dynamic optimization methods and static games encompass only fractions of aspects of those problems.
In the dissertation, we propose several models of dynamic games and dynamic optimization problems, modeling the *exploitation of common renewable resources* by taking into account various aspects of the problem:

- Many players in commons. Increasing number of players regarded as decomposition of the decision making structures. To be more specific, if we consider the same mass of individuals, decomposed into units of decreasing size: from *consumers*, through *North and South*, actual countries, regions, etc. and finally actual decision makers.

- Relation between the Nash equilibria and the social optima and ways of solving *the tragedy of the commons* by Pigovian taxation or a tax-subsidy system.

- Taking into account information: *feedback form, closed loop*, delayed information.

- Self-enforcing environmental agreements with a delay in observation of defection.

- Completing and correcting previous results in this research field or finding counterexamples to common beliefs and methodological simplifications.

In dynamic games, the strategy of a player is a function which defines his/her behaviour at each time instant in the time interval considered in the game. Therefore, calculation of both, the social optima and the Nash equilibria requires solving the dynamic optimization problems.
However, finding a Nash equilibrium in dynamic games requires solving a set of dynamic optimization problems, coupled by finding a fixed point of the resulting best response correspondence in some functional space of the profiles of strategies. Due to this coupling, the problem becomes much more complicated than the analogous dynamic optimization problems. There are quite a few results in nonzero-sum dynamic games, and if the constraints appear (which is natural in real life problems, especially resource extraction problems), then the results are very rare. Therefore, unexpected behaviour of the solution may appear (irregularity, discontinuity, the nonexistence of equilibria of a certain type, the existence of many equilibria, lack of convergence). So, we try to fill in the gaps in the simplifications of dynamic games. The dissertation also contains counterexamples to some methods and hypotheses that are regarded as correct and used to solve dynamic games.

Firstly, we consider a constrained linear quadratic dynamic game, modeling the problem of exploitation of a common renewable resource in discrete time with the infinite time horizon and with increasing number of players. So, we study a discrete time, infinite horizon, a linear quadratic dynamic game model with many players and with linear state-dependent constraints on decisions of players. In this model, players can be regarded as countries or firms. There are either finitely many players or a continuum of players. To make the model realistic, we impose the constraints on strategies. As a consequence, calculation of a feedback Nash equilibrium become complicated. The model has an obvious application in a common fishery extraction problem where the players sell their catch at a common market.

We solve the social optimum problem for n-players and for the continuum of players.
When it comes to the Nash equilibrium problem, we are only able to solve it for the continuum of players case. For n-players case, we are not able to calculate it for $n \geq 2$, only negative results can be proven: that the Nash equilibrium strategies and the value functions are not of assumed regularity with respect to the state variable and showing that presence of even a very simple and apparent constraints on strategies may result in a very complicated form of the value functions and the Nash equilibria. We return to this problem in a truncation of the game later to show the reason that even in a 2-stage truncation such a continuous solution does not exist.

While looking for a Nash equilibria, the social optima, we have also found a very simple example that may be treated as a counterexample to correctness of the undetermined coefficient method or Ansatz method, used for solving the Nash equilibrium and/or optimal control problems to the correctness of a procedure often used in dynamic game theory literature. We also calculate the enforcement of a social optimum profile by various type of *Pigouvian tax* or a tax-subsidy system, both for n-players and for the continuum of players.

Non-existence of a symmetric feedback Nash equilibrium of assumed regularity in the linear quadratic problem considered before seems to be inherited from the finite time horizon truncations of the game, so we solve a feedback Nash equilibrium problem in a very simple 2-stage, 2-player linear-quadratic dynamic game, a truncation of the model which was studied before with the infinite time horizon. As a result, we found that the presence of simple linear state-dependent constraints results in the non-existence of a continuous symmetric feedback Nash equilibria, whereas the existence of the continuum of discontinuous symmetric feedback Nash equilibria (discontinuous with respect to the state variable). Our result is counter-intuitive to the common belief in the continuity of Nash equilibria for linear-quadratic dynamic games with concave payoffs.

While previous research works deal with the specific value of the discount factor β, given by the so-called *golden rule*, now we solve the social optimum problem considered before for more general class of linear-quadratic dynamic games with only one player, called social planner and for more general β instead of *the golden rule* β. So, we consider a discrete time linear-quadratic dynamic optimization problem with linear state-dependent constraints. We solve the problem in the infinite time horizon and its finite horizon truncations. Although it seems simple in its linear quadratic form, calculation of the optimal control is nontrivial.

Next, we study a general class of dynamic optimization problems. We derive general rules stating what kind of errors in calculation or computation of the value function does not lead to errors in calculation or computation of optimal control. This general result concerns not only errors resulting from using the numerical methods but also errors resulting from some preliminary assumptions related to constraints on the value functions. The results are illustrated by a motivating example of discrete time Fish Wars model, proposed by Levhari and Mirman, with singularities in payoffs.

Finally, we study a continuous time version of the Fish Wars model with the infinite time horizon, linear state equation, and state-dependent linear constraints on controls. We calculate the social optimum and a Nash equilibrium which always leads to the depletion of the resource even if the social optimum results in its sustainability. We propose two ways of solving the problems of enforcing social optimality: either by a tax-subsidy system or by an environmental agreement even if we assume that it takes time to detect any defection of a player. We also propose a general algorithm for finding the financial incentives enforcing the socially optimal profile in a large class of differential games.