Using Game Theory for Distributed Control Engineering

Rantzer, Anders

ISSN 0280–5316
ISRN LUTFD2/TFRT--7620--SE

Using Game Theory for Distributed Control Engineering

Anders Rantzer

Automatic Control LTH
Lund University
July 2008

Department of Automatic Control
Lund University
Box 118
SE-221 00 Lund Sweden

Document name

Date of issue
July 2008

Document Number
ISRN LUTFD2/TFRT--7620--SE

Author(s)
Anders Rantzer

Supervisor

Sponsoring organisation
Swedish Foundation for Strategic Research

Title and subtitle
Using Game Theory for Distributed Control Engineering

Abstract

The purpose of this paper is to show how ideas from game theory and economics may play an important role for decentralized controller design in complex engineering systems. The focus is on coordination through prices and iterative price adjustments using a gradient method. We give a quantitative bound on adjustment rates that is sufficient to guarantee global convergence towards a Nash equilibrium of control strategies. The equilibrium is an optimal solution to the corresponding team decision problem, where a distributed set of controllers cooperate to optimize a common objective. The method is illustrated on control of a vehicle formation (e.g. automobiles on the road) where the objective is to maintain desired vehicle distances in presence of disturbances.

Key words

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title
0280–5316

ISBN

Language
English

Number of pages

Recipient's notes

Security classification

Presented at Games 2008, 3rd World Congress of the Game Theory Society

Using Game Theory for Distributed Control Engineering

Anders Rantzer

***Abstract*— The purpose of this paper is to show how ideas from game theory and economics may play an important role for decentralized controller design in complex engineering systems. The focus is on coordination through prices and iterative price adjustments using a gradient method. We give a quantitative bound on adjustment rates that is sufficient to guarantee global convergence towards a Nash equilibrium of control strategies. The equilibrium is an optimal solution to the corresponding team decision problem, where a distributed set of controllers cooperate to optimize a common objective. The method is illustrated on control of a vehicle formation (e.g. automobiles on the road) where the objective is to maintain desired vehicle distances in presence of disturbances.**

I. Introduction

How should control equipments distributed across the electricity network cooperate to find new transmission routes when a power line is broken? How should the electronic stabilization programme of an automobile use measurements from wheels and suspensions and decide how to use available brakes and engine power to recover from a dangerous situation? How should radio transmission power between cellphones and base station in a telephone network be adjusted to accommodate for a optimal use of the radio channel when the network load is high?

These are all problems of distributed control engineering, where several units need to cooperate with access to different information and with bounds on the communication between them. The classical engineering approach to such problems is to assign one controller for each task and minimize the interaction between them. However, the increasing complexity of engineering systems makes it desirable to go beyond the traditional methods and create a systematic theory for decentralized decision-making and policy updates in dynamical systems. The purpose of this paper is to show how ideas from game theory and economics may play an important role for this purpose.

We consider systems described by differential equations or difference equations. There is a set of agents, each equipped with some decision variables influencing the dynamics of the system. Every agent tries to optimize his own objective defined in terms of the system dynamics. However, the decisions by one agent will also influence the others and we seek methods to handle this interaction through iterative negotiations between the agents.

Iterative processes with provable convergence to a Nash equilibrium for general classes of games are hard to obtain. Similar difficulties appear in general equilibrium theory of economics when it comes to price negotiations aiming to reach a Walras equilibrium. However, many engineering applications can be viewed as "team decision problems", with an over-all design objective that is common to all agents [6]. This is a special class of games, which are considerably easier to analyze. Introduction of prices makes it possible to split the team problem into an equivalent game where each agent has a local objective which is "linear in money". A classical argument then shows that price iterations in the gradient direction converge towards the desired equilibrium. See [12, Example 7, page 105]. For a convex-concave function, gradient dynamics in continuous time were proved by Arrow, Hurwicz and Usawa to converge globally towards the saddle-point [1]. The gradient iteration is known as the saddle point algorithm, or Usawa's algorithm. The method to decompose a team problem has been used extensively in methods for large-scale optimization, where it is known as dual decomposition.

Distributed control problems and the relationship to team decision problems have recently gained renewed attention in the engineering literature. It has been shown that a collection of controllers with access to different sets of measurements can be designed using finite-dimensional convex optimization to act optimally as a team. The study of dynamic team problems was initiated already in 1968 by Witsenhausen [13], who also pointed out a fundamental difficulty related to information propagation. Some special types of team problems were solved in the 1970's [11], [5], but the research activity in the area remained moderate until recently. Distributed control problems with spatial invariance was exploited in [2], [3] and conditions for convexity were derived in [10], [9].

In our previous paper [7] a linear quadratic stochastic optimal control problem was solved for a state feedback control law with covariance constraints. The method gives a non-conservative extension of linear quadratic control theory to distributed problems with bounds on the rate of information propagation. An output feedback version of the problem was solved in [8] and for both finite and infinite time horizons in [4].

A. Rantzer is with Automatic Control LTH, Lund University, Box 118, SE-221 00 Lund, Sweden, rantzer at control.lth.se.

At the same time, (p_1^v, p_2^v) are the Lagrange multipliers corresponding to the two equality constraints.

Below is a formal statement of the convergence result (without the sign constraints on x_j of the original paper) [1]:

Theorem 1 (Arrow, Hurwicz, Usawa): Assume that $V \in C^1(\mathbf{R}^n)$ is strictly convex with gradient ∇V, while G and H are positive definite and R has full row rank. Then, all solutions to

$$\begin{aligned} \dot{x} &= -G\left[(\nabla V)^T - R^T p\right] \\ \dot{p} &= -HRx \end{aligned}$$

converge to the unique saddle point (x_*, p_*) attaining

$$\max_p \min_x \left[V(x) - p^T Rx\right] \tag{5}$$

Proof. Let $\phi(x,p) = V(x) - p^T Rx$. Then

$$\dot{x} = G\left[\nabla_x \phi(x,p)\right]^T \qquad \dot{p} = -H\left[\nabla_p \phi(x,p)\right]^T$$

Define the Lyapunov function

$$W(x,p) = \frac{1}{2}\left(|x - x_*|^2_{G^{-1}} + |p - p_*|^2_{H^{-1}}\right)$$

Then

$$\begin{aligned} \dot{W} &= \dot{x}^T G^{-1}(x - x_*) + \dot{p}^T H^{-1}(p - p_*) \\ &= \left[\nabla_x \phi(x,p)\right](x - x_*) - \left[\nabla_p \phi(x,p)\right](p - p_*) \\ &\leq \left[\phi(x,p) - \phi(x_*,p)\right] - \left[\phi(x,p) - \phi(x,p_*)\right] \\ &= \left[\phi(x,p_*) - \phi(x_*,p_*)\right] - \left[\phi(x_*,p) - \phi(x_*,p_*)\right] \leq 0 \end{aligned}$$

with equality if and only if $x = x_*$. Hence, by LaSalle's theorem, $(x(t), p(t))$ tends towards M, the largest invariant set in the subspace $x = x_*$. Invariance means that $\dot{x} = 0$, hence $\nabla V(x)^T = R^T p$, so the only point in M is (x_*, p_*). This completes the proof. □

To concretize the result for the vehicle formation, the response to a brief disturbance in v is plotted in Figure 2, using gradient dynamics when (1)-(3) are given by

$$\begin{aligned} &6(x_{11} - v)^2 + p_1 x_{11} \\ &3(x_{12} - x_{22})^2 - p_1 x_{12} + p_2 x_{22} \\ &2(x_{23} - x_{33})^2 + 2(x_{33})^2 - p_2 x_{23} \end{aligned}$$

Notice that the cost of the first vehicle quickly recovers, but there is a poorly damped oscillation in the response of the second and third vehicle. This reflects the fact that only the stationary equilibrium is optimized, not the transient dynamics. Hence it is natural to ask: Can the same idea of dual decomposition be used to get a distributed scheme for design of dynamic controllers? To address this question, a more abstract version of the theory will be introduced in the next few sections, before returning to the vehicle formation problem.

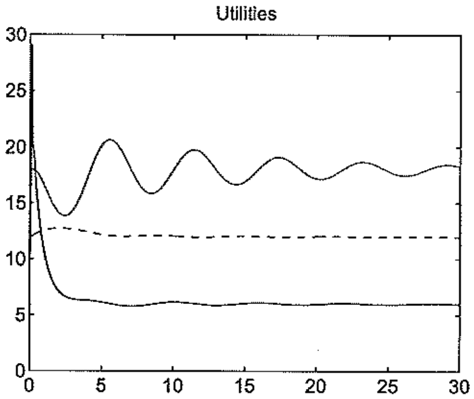

Fig. 2. The dynamics of the three cost functions when the reference value for the first vehicle is subject to a transient disturbance. The cost of the first agent quickly recovers, but there is a poorly damped oscillation in the response of the second and third vehicle.

III. Basic notions of game theory

A (strategic) *game* is defined by a map

$$\mu = (\mu_1, \ldots, \mu_J) \mapsto \left(\mathcal{V}_1(\mu), \ldots, \mathcal{V}_J(\mu)\right) \tag{6}$$

where μ_j is the *strategy* of player j and $\mathcal{V}_j(\mu) \in \mathbf{R}$ is the *payoff* for player j. The notation $(\mu_j, \widehat{\mu}_{-j})$ is used to denote set of strategies that is equal to $\widehat{\mu}$ in all entries except μ_j. A *Nash equilibrium* of the game is a set of strategies $\widehat{\mu} = (\widehat{\mu}_1, \ldots, \widehat{\mu}_J)$ such that

$$\mathcal{V}_j(\widehat{\mu}) \geq \mathcal{V}_j(\mu_j, \widehat{\mu}_{-j}) \text{ for all } \mu_j \tag{7}$$

Given $\epsilon > 0$, a *Nash ϵ-equilibrium* of the game is a set of strategies $\widehat{\mu} = (\widehat{\mu}_1, \ldots, \widehat{\mu}_J)$ such that

$$\mathcal{V}_j(\widehat{\mu}) + \epsilon \geq \mathcal{V}_j(\mu_j, \widehat{\mu}_{-j}) \text{ for all } \mu_j \tag{8}$$

A *potential game* is a game for which there exists a *potential function* $\Phi : \mu \mapsto \mathbf{R}$ such that

$$\operatorname{sgn}\left[\mathcal{V}_j(\widehat{\mu}) - \mathcal{V}_j(\mu)\right] = \operatorname{sgn}\left[\Phi(\widehat{\mu}) - \Phi(\mu)\right]$$

whenever $\mu_i = \widehat{\mu}_i$ for $i \neq j$.

For the purposes of this paper, we also introduce the following slightly more general concept: A *minmax potential game* is a game for which there exists a *potential function* $\Phi : \mu \mapsto \mathbf{R}$ and a partitioning $\{1, \ldots, J\} = \mathcal{J}_1 \cup \mathcal{J}_2$ such that

$$\operatorname{sgn}\left[\mathcal{V}_j(\widehat{\mu}) - \mathcal{V}_j(\mu)\right] = \begin{cases} \operatorname{sgn}\left[\Phi(\widehat{\mu}) - \Phi(\mu)\right] & \text{if } j \in \mathcal{J}_1 \\ -\operatorname{sgn}\left[\Phi(\widehat{\mu}) - \Phi(\mu)\right] & \text{if } j \in \mathcal{J}_2 \end{cases}$$

whenever $\mu_i = \widehat{\mu}_i$ for $i \neq j$. The players in $\mathcal{J}_1$ maximize the potential, while the players in $\mathcal{J}_2$ minimize the potential.

A *team problem* with *team payoff* $\mathcal{V}$ is a game where $\mathcal{V}_1 = \cdots = \mathcal{V}_J = \mathcal{V}$. If $\mathcal{V}(\widehat{\mu}) = \max_\mu \mathcal{V}(\mu)$, then $\widehat{\mu}$ is called an *optimal set of strategies* for the team problem. (Moreover, $\widehat{\mu}$ is a Nash equilibrium of every potential game with potential function $\mathcal{V}$.)

Theorem 3: Given $\mathcal{V}_1, \dots, \mathcal{V}_J, \mathcal{W}_1, \dots, \mathcal{W}_K$ as in (13)-(14), the corresponding game is a minmax potential game with potential $\Phi(\mu, x, \lambda)$ defined by (15). If $(\widehat{\mu}, \widehat{x}, \widehat{\lambda})$ is a Nash equilibrium of this game, then $\widehat{\mu}$ is a Nash equilibrium for the team problem defined by (11)-(12).

Moreover, suppose $\min_\lambda \max_{\mu,x} \Phi(\mu, x, \lambda)$ is attained in a point $(\widehat{\mu}, \widehat{x}, \widehat{\lambda})$ satisfying the following (local) conditions: There exists $\epsilon > 0$ such that for λ in a neighborhood of $\widehat{\lambda}$

$$\max_{\mu,x} \Phi(\mu, x, \lambda) \geq \Phi(\widehat{\mu}, \widehat{x}, \widehat{\lambda}) + \epsilon \|\lambda - \widehat{\lambda}\|^2 \tag{16}$$

and $\arg\max_{(\mu,x)} \Phi(\mu, x, \lambda) \to (\widehat{\mu}, \widehat{x})$ as $\lambda \to \widehat{\lambda}$. Then, the sequence $\{(\mu(\tau), x(\tau), \lambda(\tau))\}_{\tau=1}^{\infty}$ defined by

$$\begin{aligned}
&(\mu_j(\tau+1), x_j(\tau+1)) \\
&= \arg\max_{(\mu_j, x_j)} \left(V_j(x_j, \mu_j) + \sum_{i_k = j} \langle \lambda_k, a_k(x_j) \rangle + \sum_{j_l = j} \langle \lambda_l, b_j(x_j, \mu_j) \rangle \right)
\end{aligned}$$

$$\lambda_k(\tau+1) = \lambda_k(\tau) - \frac{1}{2}\gamma_k \left[a_k(x_{i_k}(\tau)) + b_k(x_{j_k}(\tau), \mu_{j_k}(\tau)) \right]$$

converges (globally) towards the Nash equilibrium $(\widehat{\mu}, \widehat{x}, \widehat{\lambda})$ as $\tau \to \infty$, provided that the numbers $\gamma_1, \dots, \gamma_K > 0$ are small enough to make $\max_{\mu,x} \Phi(\mu, x, \lambda) - \sum_k \|\lambda_k\|^2 / \eta_k$ a concave function of λ for some $\eta_k > \gamma_k$.

Proof. When $(\mu_i, x_i) = (\widehat{\mu}_i, \widehat{x}_i)$ for $i \neq j$, we have

$$\mathcal{V}_j(\widehat{\mu}, \widehat{x}, \widehat{\lambda}) - \mathcal{V}_j(\mu, x, \lambda) = \Phi(\widehat{\mu}, \widehat{x}, \widehat{\lambda}) - \Phi(\mu, x, \lambda)$$

Similarly, when $\lambda_l = \widehat{\lambda}_l$ for $l \neq k$

$$\mathcal{W}_k(\widehat{\mu}, \widehat{x}, \widehat{\lambda}) - \mathcal{W}_k(\mu, x, \lambda) = \Phi(\mu, x, \lambda) - \Phi(\widehat{\mu}, \widehat{x}, \widehat{\lambda})$$

Hence the conditions for a minmax potential game hold. If $(\widehat{\mu}, \widehat{x}, \widehat{\lambda})$ is a Nash equilibrium of this game, then

$$\begin{aligned}
\mathcal{V}(\mu) &\leq \sup_{(\mu_j, x_j)} \inf_{\lambda} \Phi(\mu, x, \lambda) \\
&\leq \sup_{(\mu_j, x_j)} \Phi(\mu, x, \widehat{\lambda}) = \Phi(\widehat{\mu}, \widehat{x}, \widehat{\lambda}) = \mathcal{V}(\widehat{\mu})
\end{aligned}$$

The first inequality follows directly from the definition of $\mathcal{V}(\mu)$, while the two equalities follow from the assumption that $(\widehat{\mu}, \widehat{x}, \widehat{\lambda})$ is a Nash equilibrium. Altogether, this proves that $\widehat{\mu}$ is a Nash equilibrium for the team problem.

To prove convergence towards the Nash equilibrium, introduce

$$h = \begin{bmatrix} h_1 \\ \vdots \\ h_K \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\gamma_1 \Big(a_1(x_{i_1}(\tau)) + b_1(x_{j_1}(\tau), \mu_{j_1}(\tau)) \Big) \\ \vdots \\ \frac{1}{2}\gamma_K \Big(a_K(x_{i_K}(\tau)) + b_K(x_{j_K}(\tau), \mu_{j_K}(\tau)) \Big) \end{bmatrix}$$

$$\langle \lambda, \eta^{-1}\lambda \rangle = \sum_k \langle \lambda_k, \lambda_k \rangle / \eta_k$$

We also use the notation $\lambda^+ = \lambda(\tau+1)$, $x^+ = x(\tau+1)$ and $\mu^+ = \mu(\tau+1)$. Let $U(\lambda) = \max_{\mu,x} \Phi(\mu, x, \lambda)$. Then

$$\begin{aligned}
U(\lambda) - U(\widehat{\lambda}) &= \Phi(\mu^+, x^+, \lambda) - \max_{\mu,x} \Phi(\mu, x, \widehat{\lambda}) \\
&\leq \Phi(\mu^+, x^+, \lambda) - \Phi(\mu^+, x^+, \widehat{\lambda}) \\
&= \langle \lambda - \widehat{\lambda}, a(x^+) + b(x^+, \mu^+) \rangle \\
&= \langle \lambda - \widehat{\lambda}, 2\gamma^{-1}h \rangle
\end{aligned} \tag{17}$$

The condition (16) can be written

$$U(\lambda) \geq U(\widehat{\lambda}) + \epsilon \|\lambda - \widehat{\lambda}\|^2$$

for λ close to $\widehat{\lambda}$. However, U is convex by construction, so

$$U(\lambda) \geq U(\widehat{\lambda}) + \|\lambda - \widehat{\lambda}\| \min\{\epsilon \|\lambda - \widehat{\lambda}\|, \mu\}$$

must hold globally for sufficiently small $\mu > 0$. Combining this with (17) gives

$$\|\lambda - \widehat{\lambda}\| \min\{\epsilon \|\lambda - \widehat{\lambda}\|, \mu\} \leq \langle \lambda - \widehat{\lambda}, 2\gamma^{-1}h \rangle$$

thus

$$\min\{\epsilon \|\lambda - \widehat{\lambda}\|, \mu\} \leq 2\|\gamma^{-1}h\|$$

To prove that $\lim_{\tau\to\infty} \lambda(\tau) = \widehat{\lambda}$, it therefore remains to prove that $\lim_{\tau\to\infty} \|h(\tau)\| = 0$. Concavity of $U(\lambda) - \langle \lambda, \eta^{-1}\lambda \rangle$ gives

$$\begin{aligned}
&U(\lambda + h) - \langle \lambda + h, \eta^{-1}(\lambda + h) \rangle \\
&\quad + U(\lambda - h) - \langle \lambda - h, \eta^{-1}(\lambda - h) \rangle \\
&\leq 2U(\lambda) - 2\langle \lambda, \eta^{-1}\lambda \rangle
\end{aligned}$$

for $h \neq 0$. Hence

$$\begin{aligned}
U(\lambda^+) &= U(\lambda - h) \\
&\leq 2U(\lambda) - U(\lambda + h) + 2\langle h, \eta^{-1}h \rangle \\
&\leq 2U(\lambda) - \Phi(\mu^+, x^+, \lambda + h) + 2\langle h, \eta^{-1}h \rangle \\
&= U(\lambda) - \langle h, a(x^+) + b(x^+, \mu^+) \rangle + 2\langle h, \eta^{-1}h \rangle \\
&= U(\lambda) + 2(\eta^{-1} - \gamma^{-1})\|h\|^2
\end{aligned}$$

In particular

$$\begin{aligned}
U(\widehat{\lambda}) &\leq U(\lambda(T)) \\
&= \sum_{\tau=0}^{T-1} \left[U(\lambda(\tau+1)) - U(\lambda(\tau)) \right] + U(\lambda(0)) \\
&\leq 2(\eta^{-1} - \gamma^{-1}) \sum_{\tau=0}^{T-1} \|h(\tau)\|^2 + U(\lambda(0))
\end{aligned}$$

so $\sum_{\tau=0}^{\infty} \|h(\tau)\|^2$ is bounded and $\lim_{\tau\to\infty} \|h(\tau)\| = 0$. This proves that $\lim_{\tau\to\infty} \lambda(\tau) = \widehat{\lambda}$. By assumption, this also gives $\lim_{\tau\to\infty}(\mu(\tau), x(\tau)) = (\widehat{\mu}, \widehat{x})$, so the proof is complete. □

VI. Vehicle formations reconsidered

For optimization of dynamic controllers in the vehicle formation, it is useful to quantify the resulting stationary dynamics when v is a given discrete time stochastic process. For an example of stable (but sub-optimal) dynamics, one could let each vehicle optimize its position based on prices determined by the following discrete time version of the gradient dynamics studied before:

$$\begin{aligned}
p_1(t+1) &= p_1(t) + h_1 \left[x_{11}(t) - x_{12}(t) \right] \\
p_2(t+1) &= p_2(t) + h_2 \left[x_{22}(t) - x_{23}(t) \right]
\end{aligned} \tag{18}$$

For small adjustment rates, the discrete time dynamics is similar to the continuous time behavior, hence stable. Suppose the disturbance $v(t)$ is a Gaussian discrete time stationary stochastic process given by

$$v(t+1) = av(t) + w(t)$$

where $w(t)$ is zero mean white noise with unit variance. Figure 3 shows a simulation with $a = 0.9$ and $g_{11} = g_{12} = g_{22} = g_{23} = g_{33} = h_1 = h_1 = 0.1$. Due to the structure of the gradient dynamics, the effect of $w(t)$ will propagate through the vehicle formation, affecting the first vehicle position x_{11} at time $t+1$, the price p_1 and the second vehicle position at time $t+2$ and finally the price p_2 and the third vehicle at time $t+3$. We will now consider iterative improvement of