European Journal of Operational Research 000 (2018) 1–22

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

A linear-quadratic Gaussian approach to dynamic information acquisition[☆]

Thomas A. Weber[a,*], Viet Anh Nguyen[b]

[a] Chair of Operations, Economics and Strategy, École Polytechnique Fédérale de Lausanne, Station 5, Lausanne CH-1015, Switzerland
[b] École Polytechnique Fédérale de Lausanne, Station 5, Lausanne CH-1015, Switzerland

ARTICLE INFO

Article history:
Received 15 February 2017
Accepted 2 March 2018
Available online xxx

Keywords:
Dynamic emissions regulation
Information acquisition
Infinite-horizon optimal control
Linear-quadratic systems
Markov decision problems

ABSTRACT

We consider optimal information acquisition for the control of linear discrete-time random systems with noisy observations and apply the findings to the problem of dynamically implementing emissions-reduction targets. The optimal policy, which is provided in closed form, depends on a single composite parameter which determines the criticality of the system. For subcritical systems, it is optimal to perform "noise leveling," that is, to reduce the variance of the state uncertainty to an optimal level and keep it constant by a steady feed of information updates. For critical systems, the optimal policy is "noise attenuation," that is, to substantially decrease the variance once and never acquire information thereafter. Finally for supercritical systems, information acquisition is never in the best interest of the decision maker. In each case, an explicit expression of the value function is obtained. The criticality of the system, and therefore the tradeoff between spending resources on the control or on information to improve the control, is influenced by a "policy parameter" which determines the importance a decision maker places on uncertainty reduction. The dependence of the system performance on the policy parameter is illustrated using a practical climate-control problem where a regulator imposes state-contingent taxes to probabilistically attain emissions targets.

1. Introduction

The effective control of a stochastic system critically depends on sufficient information about its state. The optimal acquisition of state information balances the expected increase of the decision maker's value with the cost of the signal that is being acquired. The quality of the state information determines the precision with which, at any given point in time, the decision maker can condition the choice of the best available action on the actual system behavior. For example, when trying to implement greenhouse-gas emission-reduction targets a regulator can set taxes or quotas. The target for such stock pollutants are usually expressed in terms of aggregate emissions that should stay within a carbon budget. The latter is almost linearly related to the projected increase in average temperature (IPCC, 2014). Hence, while it is possible to steer aggregate emissions *in expectation* to a given target by dynamically setting carbon prices, the probability of the *actual* state being close to the target hinges on the quality of the acquired information about the emissions level. In this paper, we provide a closed-form solution to the combined control and information-acquisition problem for linear systems and quadratic costs with one-dimensional state. The application to emissions control is then discussed based on an established model by Hoel and Karp (2002) using recent global emissions data and targets (IEA, 2015).

The importance of combining the optimal control of a system with the estimation of its state was first recognized for engineering applications (Meier, 1965). Upon investigation, it was quickly realized that the estimation and optimization problems can be decoupled, in both discrete time (Striebel, 1965) and continuous time (Wonham, 1968), resulting in a "separation principle" (Davis, 1977; Fleming & Rishel, 1975). Yet, in virtually all of the extant work, the precision of the information about the state is taken as given.

☆ The authors would like to thank two anonymous referees and participants of the Moscow Summer Academy on Economic Growth and Governance of Natural Resources at Moscow State University (Russia), the 12th International Conference on Operations Research in Havana (Cuba), the 10th Conference of the Association of Asia-Pacific Operational Research Societies (APORS) in Kuching (Malaysia), the 2015 INFORMS Annual Meeting in Philadelphia, PA (USA), and the 10th Conference of the International Federation of Operational Research Societies (IFORS) in Barcelona (Spain) for their helpful comments and suggestions.

* Corresponding author.
E-mail addresses: thomas.weber@epfl.ch (T.A. Weber), viet-anh.nguyen@epfl.ch (V.A. Nguyen).

https://doi.org/10.1016/j.ejor.2018.03.003

Please cite this article as: T.A. Weber, V.A. Nguyen, A linear-quadratic Gaussian approach to dynamic information acquisition, European Journal of Operational Research (2018), https://doi.org/10.1016/j.ejor.2018.03.003

That is, the decision maker remains unconcerned with the problem of acquiring an appropriate amount of information.[1] Here we consider the linear-quadratic control problem with costly information of varying precision as an archetypical case which can be solved completely. We show that the separation principle applies and that the best information-acquisition policy achieves an optimal noise level over and above the system noise.[2] The optimal information-acquisition policy is implemented by a threshold policy that at any time depends on the current variance of the decision maker's beliefs about the state of the system. The precise nature of this policy can be fully characterized by a "discriminant" of the problem.

In practice, a decision maker's incentives to acquire information involve more complex considerations than what is contained in the initially formulated combined control and information-acquisition problem. For example, it may be desirable to maximize the probability of the system's state to be close to a specified target state at a given date. For this, the decision maker can tune a "policy parameter" which describes the tradeoff between the control problem (referred to as *system-stabilization problem*) and the uncertainty-reduction problem (referred to as *information-acquisition problem*). The policy parameter modulates therefore the decision maker's information-acquisition effort. Our analysis addresses the comparative statics of the problem with respect to the policy parameter, thus illustrating the structural insights that can be obtained from a closed-form solution to the decision problem.

1.1. Literature

Blackwell (1951) showed that information is generally beneficial to decision makers and that more informative sources of information are of (weakly) greater value. Conversely, a pure increase in uncertainty about the state by adding noise to an information source corresponds to a "garbling" and therefore must (weakly) decrease the information value (DeGroot, 1962). The finding carries over to a Bayesian setting (Kihlstrom, 1984), and the decision maker's value for state observations with Gaussian noise is decreasing in the variance of those information sources. While there is a fairly rich work on the value of information in a (quasi-)static setting (LaValle, 1968; Lawrence, 1999), the literature on the sequential acquisition of information, after a promising start, experienced a long hiatus.[3] The seminal contribution by Wald (1947) provides a general approach to the information-acquisition problem, and then concentrates efforts on examining distributionally robust experimentation to improve a statistical decision function. Building on pieces of that initial framework, Moscarini and Smith (2001) perform an interesting analysis of (nonrobust) information acquisition in continuous time, by controlling the diffusion of a Brownian motion through the continuous acquisition of a somewhat peculiar, specially adapted sampling process so as to inform a binary decision. McCardle (1984) considers information acquisition in a discrete-time dynamic setting, where a firm gathers information to reduce uncertainty about a technology-adoption decision. The latter leads to an optimal stopping problem, where at an upper belief threshold the firm decides to adopt the technology (and stop information acquisition), at a lower belief threshold to reject the technology (and stop information acquisition), and otherwise to continue gathering information. The underlying problem of optimally stopping in a Markovian setting with costly information about an imperfectly observable state was discussed by Monahan (1980), and results about the convexity of policy regions are summarized by Lovejoy (1987). Similarly, Moore and Whinston (1986, 1987) discuss sequential information acquisition, followed by a final action.[4] By contrast, we are concerned here with problems where control interventions and information acquisition coexist from period to period. For example, in operations management, costly information about demand can help improve inventory-management decisions (DeCroix & Mookerjee, 1997). In a newsvendor setting with independent and identically distributed consumers, Milgrom and Roberts (1988) find that it is best to either survey none or all of them, i.e., to acquire either no or full information; the reason for this is a convexity in the value of additional information. In contrast to this, Fu and Zhu (2010), by using forecast-aggregation techniques, obtain a concave information value which generically leads to the optimality of intermediate levels of information acquisition. In a linear-quadratic setting, Bansal and Başar (1989) take an information-theoretic approach separating the measurement (or communications) task from the control task, in an iterative discrete-time framework. The authors consider a similar setting in continuous time (Başar & Bansal, 1994); see also Yüksel and Başar (2013, Ch. 11) for a summary of this decentralized encoder-decoder approach with ample additional references. Sims (2003) limits the flow of information by imposing a bound on the Shannon channel capacity, which then makes the amount of information collected over time subject to optimization. Provided the channel capacity is not too low, the best policy approaches an optimal signal-to-noise ratio in the long run, thus closing in on a stationary variance of the state given a stationary variance in the observational noise. Here we also consider a linear-quadratic problem setup, yet instead of imposing an exogenous limit on the information flow, we allow for a linear cost of the precision of the state observation.

The standard linear-quadratic Gaussian (LQG) optimal control problem consists in choosing the input for a linear system so as to maximize the expectation of a quadratic functional which depends on the realized trajectories of the state, the output, and the control (Anderson & Moore, 1971; Athans, 1972a). The linear-quadratic setup appears naturally in many managerial and policy-relevant contexts, such as inventory control (Holt, Modigliani, Muth, & Simon, 1960; West, 1986), error-correction mechanisms (Salmon, 1982), production smoothing and scheduling (Gallego, 1990; Naish, 1994), dynamic oligopoly (Kydland, 1975; Fudenberg and Tirole, 1986), monetary and fiscal policy (Benigno & Woodford, 2004; Pindyck & Roberts, 1974), forecasting of economic equilibria (Townsend, 1983), nonlinear pricing with learning (Bonatti, 2011), and dynamic regulation (Auray, Mariotti, & Moizeau, 2011; Friedman, 1981), to just name a few. The optimal value of the objective depends on the quality of the state observations. The corresponding linear-quadratic combined estimation and control problem was solved by Kalman (1960), resembling results by Thiele (1880) (see Lauritzen, 1981). Athans (1972b) provides an algorithm for optimally switching among a finite number of sensors in continuous time, effectively solving an offline sensor-selection

[1] A notable exception is the dynamic sensor-selection problem introduced by Athans (1972b) which, from an engineering standpoint, amounts to scheduling the best camera view onto the state. Due to the combinatorial nature of this problem, it leads only to an algorithmic solution without much structural insight; see also Section 1.1.

[2] The findings can be interpreted in terms of signal-to-noise ratios, but they appear most naturally in terms of the additionally tolerated observational noise.

[3] There is significant work on the exploration-versus-exploitation tradeoff, inherent in the multi-armed bandit problem, but here information acquisition is mixed with reward-oriented actions and they are difficult to disentangle; see Gittins, Glazebrook, and Weber (2011) for details.

[4] Applications in healthcare also have this feature, where a sequence of tests ("screening" actions) may be followed by a treatment; this is complicated by the fact that the disease progression or population characteristics may be nonstationary. Tsodikov and Yakovlev (1991) examine aperiodic cancer screening, while Maillart, Ivy, Ransom, and Diehl (2008) use a hidden Markov-chain approach. The latter is also used by Cipriano and Weber (2018) to determine when and how much sampling is needed before the decision to discontinue a public health screening for a population with declining hepatitis C prevalence.