OPTIMAL SELECTION OF INFORMATION WITH RESTRICTED STORAGE CAPACITY

L. Pronzato

Laboratoire I3S, CNRS–UPRESA 6070
Bât. 4, 250 rue A. Einstein,
Sophia Antipolis, 06560 Valbonne, France

ABSTRACT

We consider the situation where n items have to be selected among a series of N presented sequentially, the information contained in each item being random. The problem is to get a collection of n items with maximal information. We consider the case where the information is additive, and thus need to maximize the sum of n independently identically distributed random variables x_k observed sequentially in a sequence of length N. This is a stochastic dynamic–programming problem, the optimal solution of which is derived when the distribution of the x_k's is known. The asymptotic behaviour of this optimal solution (when N tends to infinity with n fixed) is considered. A (forced) certainty–equivalence policy is proposed for the case where the distribution is unknown and estimated on–line.

1. INTRODUCTION AND PROBLEM STATEMENT

We consider the situation where a sequence of items y_k is proposed, e.g. for further delayed processing, the information contained in each y_k being measured by a scalar variable x_k. The sequence $\{y_k\}$ has length N and the storage capacity is $n < N$. The problem is then to derive on–line a decision rule for the storage of the y_k's which maximizes the information contained in the items. The storage is permanent: items are selected forever, that is any item selected at time j cannot be replaced by another one at time $k > j$.

Another possible setup is when a scalar parameter θ has to be estimated from observations y_k, described by $y_k = f(\theta, z_k) + \epsilon_k$ with $\{\epsilon_k\}$ an i.i.d. sequence of random variables. N experiments, characterized by the z_k's are proposed, but only n can be performed. Their selection forms a problem of *sequential experiment design*. The criterion is Fisher information, given by $\sum_{i=1}^{n} \omega^2(\theta^0, z_{k_i})$, with $k_i \in \{1, \ldots, N\}$, $\omega(\theta, z) = \partial f(\theta, z)/\partial\theta$, and θ^0 some prior nominal value of θ.

We shall assume throughout the paper that the x_k's are independent random variables. The dependent case (e.g. Markov process), and the case where information is not additive (e.g. when $y_k = f(\boldsymbol{\theta}, z_k) + \epsilon_k$, with $\boldsymbol{\theta}$ a p-dimensional vector), deserve further studies.

Let $\{u_k\}$ be the decision sequence, with $u_k = 1$ if y_k is stored/observed at time k, and $u_k = 0$ otherwise. The problem is then to maximize

$$J_N(\boldsymbol{u}_1^N) = \sum_{k=1}^{N} u_k x_k \,, \tag{1}$$

with $\boldsymbol{u}_1^N = (u_1, \ldots, u_N)$ satisfying the constraint

$$\sum_{k=1}^{N} u_k \leq n \,, \tag{2}$$

the constraint will be saturated at the optimum, see (3). For any sequence $\boldsymbol{u}_1^N$ and any time (or step) j, $1 \leq j \leq N$, let a_j be the number of items already stored, that is $a_j = \sum_{k=1}^{j-1} u_k$ with $a_1 = 0$, and $c(j, a_j, \boldsymbol{u}_1^N)$ be the gain–to–go,

$$c(j, a_j, \boldsymbol{u}_j^N) = \sum_{k=j}^{N} u_k x_k = x_j u_j + c(j+1, a_{j+1}, \boldsymbol{u}_{j+1}^N) \,.$$

Note that x_j is known when u_j has to be chosen. Two on–line decision policies are considered in Section 2, which aim at maximizing the gain–to–go. Their asymptotic behaviour ($N \to \infty$) is studied in Section 3. Special attention is devoted to the *optimal* (closed–loop) policy. The case where the distribution of the x_k's is unknown is considered in Section 4.

2. DECISION POLICIES

The x_k's are assumed to be independent non–negative random variables with probability measure $\mu(\cdot)$. We shall denote $F(\cdot)$ the c.d.f. of x, $E\{.\}$ the expectation