Asymptotically Optimal Decisions for Sequential Screening and Resource Allocation

Luc Pronzato

Laboratoire I3S, CNRS–UNSA, bât. Euclide,
Les Algorithmes, 2000 route des Lucioles,
BP 121, 06903 Sophia-Antipolis Cedex, France
(email: pronzato@unice.fr)

Abstract. We consider the problem of maximizing the expected sum of n variables X_k chosen sequentially in an i.i.d. sequence of length N, $N \to \infty$, which is equivalent to a resource allocation problem. When the distribution of the X_k's is known and N is fixed, the optimal threshold for accepting X_k is given by a backward recurrence equation. A simple open-loop solution is suggested, which is asymptotically optimal ($N \to \infty$ with n fixed) for a large class of distributions and is easily implementable when the distribution of the X_k's is unknown and estimated on-line.
Keywords. Resource Allocation, Sequential Design, Secretary Problem, Asymptotic Rule.

1 Introduction

Consider a problem of *sequential resource allocation*, where N jobs arrive sequentially, with X_k the value of job k. There are N machines available, and the i-th machine performs the job successfully with probability p_i. The relative ranks of the p_i's are known. The problem is to allocate machines to arriving jobs, so as to maximize the total expected reward, that is, $E\{\sum_{k=1}^{N} p_{j_k} X_k\}$, where the random variable j_k labels the machine allocated to job k. (If there are $n > N$ machines available, only the N machines with largest p_i's have to be considered, whereas if $n < N$ one can consider $N - n$ virtual machines with associated p_i equal to zero; therefore one can always assume that $n = N$). From the results in Derman et al. (1972), this assignment problem is equivalent to *sequential screening*, where one wishes to sift out the n best items among N in terms of their X value, that is, to maximize $E\{\sum_{i=1}^{n} X_{k_i}\}$, with $k_i \in \{1, \ldots, N\}$ the index of the i-th item selected. Here, the decision to select the k-th item or not must be taken *on-line*, that is, at time k, and $X_k \in \mathcal{X}$ is observed before one makes a decision. The X_k's are assumed to be independently identically distributed (i.i.d.) random variables with probability measure $\mu(\cdot)$. One can notice that this corresponds to a modification of the standard "*secretary problem*", see Ferguson (1989), Freeman (1989), Kadison (1994).

The optimal (closed-loop) solution is considered in Section 2. A simpler open-loop solution, asymptotically optimal ($N \to \infty$ with n fixed) for a large class of distributions, is suggested in Section 3. The case where the distribution of the X_k's is unknown and estimated on-line is considered in Section 4, with an approach based on *certainty equivalence*. An illustrative example is presented.

2 The Closed-Loop optimal solution

Let $\{u_k\}_k$ be the decision sequence, with $u_k = 1$ if item k is selected at step k, and $u_k = 0$ otherwise. The sequential screening problem is then to maximize $EJ_N(\boldsymbol{u}_1^N) = E\{\sum_{k=1}^N u_k X_k\}$, with $\boldsymbol{u}_1^N = (u_1, \ldots, u_N)$ satisfying the constraints $u_j \in \mathcal{U}_j = \{0,1\}\,,\ j = 1,\ldots,N\,,\ \ \sum_{k=1}^N u_k = n$, the expectation being with respect to the product measure $\mu^{\otimes N}(\cdot)$ of $(X_1, \ldots, X_N)$, and X_k being known when u_k is selected. The measure $\mu(\cdot)$ is assumed to be non degenerate (it has several support points if it is discrete) and such that $E\{|X|\} < \infty$, with $E\{.\}$ the expectation with respect to X. For any sequence $\boldsymbol{u}_1^N$ and any step j, $1 \leq j \leq N$, let a_j denote the number of items already stored; that is, $a_j = \sum_{k=1}^{j-1} u_k$, with $a_1 = 0$. The problem then reduces to an usual discrete-time stochastic control problem, where j represents time, $(a_j, X_j) \in \{0,\ldots,n\} \times \mathcal{X}$ represents the state at time j and $u_j \in \mathcal{U}_j \subseteq \{0,1\}$ the control at time j. For each $j \in \{1,\ldots,N\}$ and each policy $\boldsymbol{u}_j^N$, one can consider the conditional expected *gain-to-go* at time j given the state (a_j, X_j): $C(j, a_j, X_j, \boldsymbol{u}_j^N) = E\{\sum_{k=j}^N u_k X_k | a_j, X_j\}$. The problem then corresponds to the maximization of $C(1, a_1 = 0, X_1, \boldsymbol{u}_1^N)$, where, for all j, the control function u_j only depends on the past states $(a_1, X_1, \ldots, a_j, X_j)$. Let $\tilde{C}(j, a_j, X_j)$ denote the optimal conditional expected gain-to-go when $\boldsymbol{u}_j^N$ is chosen optimally, and $\tilde{c}(j, a_j)$ denote its expected value with respect to X_j. One can easily show (Pronzato, 2001) that it satisfies the backward recurrence equation: $\tilde{c}(j, a_j) = E\{\max[X + \tilde{c}(j+1, a_j+1), \tilde{c}(j+1, a_j)]\}$, with $\tilde{c}(j, j+n-N-1) = (N-j+1)E\{X\}$ and $\tilde{C}(j, n, X_j) = 0$ for all X_j. Having to solve a backward recurrence, and being interested in asymptotic behaviours, it is useful to reverse the course of time. Denote the optimal expected gain-to-go $\tilde{c}(j, a_j)$ by $\tilde{\lambda}_m^k$, with $k = n - a_j$ and $m = N + 1 - j - n + a_j$. Define $h(\cdot)$ by $h(\cdot) : s \mapsto h(s) = E\{(X-s)^+\}$, where $x^+ = \max(x, 0)$, so that $h(s) + s = E\{\max(X, s)\}$ and $h(M') + M' = E\{X\}$, with $M' \in \mathbb{R} \cup \{-\infty\}$ the infimum of the support of $\mu(\cdot)$. The backward recurrence above can be written $\tilde{\lambda}_m^k = \tilde{\lambda}_m^{k-1} + E\{\max(X, \tilde{\lambda}_{m-1}^k - \tilde{\lambda}_m^{k-1})\}$, and thus $\tilde{\lambda}_m^k - \tilde{\lambda}_m^{k-1} = E\{\max(X, \tilde{s}_m^k)\} = \tilde{s}_m^k + h(\tilde{s}_m^k)$, with $\tilde{s}_m^k$ the optimal threshold: $\tilde{s}_m^k = \tilde{\lambda}_{m-1}^k - \tilde{\lambda}_m^{k-1}$ for $k + m \geq 1$ and $\tilde{s}_m^0 = \infty$, $\tilde{s}_0^k = M'$ for $k \geq 1$, $m \geq 1$. This gives for any $k \geq 1$, $m \geq 1$: $\tilde{s}_m^k = \tilde{s}_{m-1}^k + h(\tilde{s}_{m-1}^k) - h(\tilde{s}_m^{k-1})$.

3 An asymptotically optimal open-loop rule

We assume that $\mu(\cdot)$ is absolutely continuous with respect to the Lebesgue measure, with density $\varphi(\cdot)$ having unbounded support ($M = \infty$). We denote

$\bar{F}(s) = 1 - F(s) = \Pr[X > s]$ and make the following assumption on the tail of $\mu(\cdot)$.
H1: $\bar{F}(\cdot)$ is twice differentiable, the density $\varphi(\cdot)$ is such that $\varphi(s) > 0$ and its derivative $\varphi'(\cdot)$ satisfies $\varphi'(s) < 0$ for s larger than some s_1. Moreover, $\bar{F}(\cdot)$ has the representation $\bar{F}(s) = \bar{F}(s_0)\exp\left[-\int_{s_0}^{s}[1/a(t)]dt\right]$, $s \geq s_0$, where the auxiliary function $a(t) > 0$ is absolutely continuous w.r.t. Lebesgue measure, with derivative $a'(t)$ having limit $\lim_{t\to\infty} a'(t) = \bar{a} \in [0, 1)$. We define $A(s) = h(s)\varphi(s)/[\bar{F}(s)]^2$ and $\bar{A} = \lim_{s\to\infty} A(s)$.

We shall use the following definitions to characterize the asymptotic performance of a given decision rule.

Definition 1 *Consider a suboptimal decision rule with expected gain-to-go λ_m^k. We shall say that it is asymptotically equivalent to the optimal rule (AEOR) if $\lambda_m^k/\tilde{\lambda}_m^k \to 1$ as $m \to \infty$, and that it is asymptotically optimal (AO) if $\tilde{\lambda}_m^k - \lambda_m^k \to 0$ as $m \to \infty$.*

We are interested in the situation when $m \to \infty$ with k fixed, that is, in the setup of Section 1 and 2, when $N \to \infty$ with n fixed. Albright and Derman (1972) consider the case where $n = N - [N\pi]$, $\pi \in (0, 1)$, $N \to \infty$, where $[x]$ rounds x towards next integer. They show that for $N \to \infty$, $\bar{F}(\tilde{s}_{[N\pi]}^{N-[N\pi]}) \to 1 - \pi$ and that the open-loop rule with constant threshold $\bar{F}^{-1}(1 - \pi)$ is AEOR. The case where k is fixed is more difficult, in the sense that, in general, constant thresholds do not provide asymptotic equivalence to the optimal rule. When the threshold is suitably chosen as a function of m and k, we have the following result, see Pronzato (2001).

Theorem 1 **(i)** $k = 1$. *If $F(\cdot)$ satisfies* **H1**, *an open-loop rule that uses thresholds $\bar{s}_m^1$ such that $m\bar{A}\bar{F}(\bar{s}_m^1) \to 1$ as $m \to \infty$ is AEOR when* $\liminf_{s\to\infty} a(s) > c > 0$ *and AO when* $\limsup_{s\to\infty} a(s) < C < \infty$. **(ii)** $k \geq 1$. *If $F(\cdot)$ satisfies* **H1** *with $\bar{a} = 0$ ($F(\cdot)$ is a von Mises function), an open-loop rule that uses thresholds $\bar{s}_m^k$ such that $m\bar{F}(\bar{s}_m^k) \to k$ as $m \to \infty$ is AEOR when* $\liminf_{s\to\infty} a(s) > c > 0$ *and AO when $a(s) \to 0$ as $s \to \infty$.*

In particular, the open-loop decision rule defined by

$$u_j(a_j, X_j) = \begin{cases} 0 & \text{if } \bar{F}(X_j) \geq \frac{n-a_j}{\bar{A}(N-j+1)+(1-\bar{A})(n-a_j)-\epsilon}\,, \\ 1 & \text{otherwise,} \end{cases} \tag{1}$$

with $0 < \epsilon << 1$, satisfies the conditions on $\bar{s}_m^k$ in Theorem 1 and gives $u_j(n, X_j) = 0$ and $u_j(a_j, X_j) = 1$ if $n - a_j \geq N - j + 1$, for any X_j. Several examples in Pronzato (2001) show that its behaviour is very close to that of the optimal strategy.

In practice, the true distribution of the X_k's is generally unknown and is estimated on-line. This situation is considered in the next section.

4 On-line estimation of the distribution

Assume now that $\mu(\cdot)$ is unknown. It is shown in Albright (1977) that the sequential resource allocation and sequential screening problems remain equivalent when $\mu(\cdot)$ belongs to a parametric family, say $\mu = \mu_\theta$, the parameters θ having a prior distribution and being updated according to Bayes rule as successive X_k's are observed. The optimal solution (closed-loop) is derived for some special families of distributions (normal, uniform, Gamma). We adopt here a simpler approach, and consider decisions based on certainty equivalence[1] (CE), see Bertsekas (1995), vol. 1, p. 248: at step j, the optimal and asymptotically optimal strategies are determined as previously, with $\bar{F}(\cdot)$ replaced by $\hat{\bar{F}}_j(\cdot)$, the empirical distribution function of X based on $X_1, \ldots, X_j$. In the CE version of the open-loop rule (1), this means that we reject X_1 and then substitute $\hat{\bar{F}}_j(\cdot)$ for $\bar{F}(\cdot)$ at any step $j > 1$. This CE open-loop rule can then be expressed directly in terms of the order statistics $\{X^*_{i,j-1}\}$:

$$u_j(a_j, X_j) = \begin{cases} 0 & \text{if } X_j < X^*_{l_j,j-1}\,, \\ 1 & \text{otherwise,} \end{cases} \tag{2}$$

with, for any j, $X^*_{l,j-1} = -\infty$ for $l > j-1$, $X^*_{0,j-1} = \infty$ and

$$l_j = \left\lceil \frac{j(n-a_j)}{\bar{A}(N-j+1) + (1-\bar{A})(n-a_j) - \epsilon} \right\rceil, \tag{3}$$

where $\lceil x \rceil$ rounds x to the nearest larger integer. Note that $l_j \geq 1$ for any $j \geq 2$, so that the first item larger than X_1 is always accepted if (2,3) are used starting at $j = 2$. Small values of X might thus be accepted when X_1 is small, hence the importance of forcing the rejection of first items. We adopt here the heuristic rule: reject X_j if $j[1 + \log(n - a_j)] \leq r^* - 1$, where $r^* = \min\{r \text{ such that } c_{r-1} \geq 1 > c_r\}$, with $c_r = 1/r + 1/(r+1) + \cdots + 1/(N-1)$, $r \leq N-1$. (For $n = 1$, this corresponds to the optimal solution of the standard secretary problem, see Lindley (1961): by rejecting the first $r^* - 1$ observations and then accepting the first X_k larger than all previous ones, one maximizes the probability of selecting the largest X.)
Example: Assume that $\mu(\cdot)$ has the density $\varphi(x) = \sqrt{2}\exp[-x^2/(2\sigma^2)]/(\sigma\sqrt{\pi})$, $x \geq 0$, with $\sigma = 1$. We take $N = 50$, which gives $r^* = 19$, and compare four different strategies:

[1] A particular feature of the problem in this context is that decisions are *neutral* with respect to estimation (see (Fel'dbaum, 1965; Patchell and Jacobs, 1971) for a definition of neutrality in control problems): the sequence $\{X_k\}_k$ is observed whatever the decision sequence $\{u_k\}_k$, and decisions have no effect on the accuracy of the estimation of the distribution. The results in Albright (1977) therefore show that *neutrality is not sufficient for certainty equivalence*, which contradicts the suggestion in Patchell and Jacobs (1971) (see also (Bar-Shalom and Tse, 1974)).

• S_1 is the CE open-loop rule defined by (2-3) with $\bar{A} = 1$, $\epsilon = 0.01$; it is suboptimal because of CE and because of the open-loop decisions;
• S_2 rejects all X_j's for $j \leq r^* - 1$ and then accepts any X_j if larger than all previous ones or if $a_j + N - j + 1 \leq n$ (when $n = 1$, it maximizes the probability of selecting the largest X_k);
• S_3 is optimal for μ;
• S_4 is the optimal-CE rule based on the empirical measure; it is suboptimal because of CE.

To compare the performances of the six strategies above, we performed 1000 independent repetitions of the sequential screening experiment. We define n^* as the number of X_k selected that are actually among the n largest values, that is, those that satisfy $X_k \geq X^*_{n,N}$. Table 1 gives the empirical means and standard deviations of $J = \sum u_k X_k$ and n^* for the six strategies, when $n = 1, 10$. The empirical means and standard deviations of the computing times T are also indicated (the computations are performed in MATLAB on a personal computer equipped with a Pentium processor 150 MHz).

		$E\{J\}$	std$\{J\}$	$E\{n^*\}$	std$\{n^*\}$	$E\{T\}$	std$\{T\}$
$n = 1$	S_1	2.02	0.55	0.291	0.45	0.07	0.025
	S_2	1.73	0.90	0.374	0.48	0.05	0.02
	S_3	2.29	0.51	0.530	0.50	0.33	0
	S_4	2.03	0.54	0.267	0.44	1.40	0.15
$n = 10$	S_1	15.61	1.90	7.24	1.18	0.09	0.03
	S_2	9.09	2.24	2.58	1.30	0.05	0.02
	S_3	16.63	1.90	8.51	0.94	2.31	0
	S_4	15.78	1.91	7.42	1.15	6.4	0.5

Table 1. Empirical means and standard-deviations of J, n^* and computing time T (in s) for strategies S_1 to S_6 ($N = 50$, 1000 repetitions)

The upper bound on J obtained for the non-sequential selection of $X^*_{1,50}$ (resp. of $X^*_{1,50}, \ldots, X^*_{10,50}$) is approximately equal to 2.51 (resp. 17.3). The decrease of performance due to the on-line character of the decisions is thus rather small for strategies S_1, S_3 and S_4. The loss due to CE is small: compare S_4 to S_3. Using open-loop decisions has also a small effect when the threshold is adaptive: compare S_1 to S_4. S_2 performs rather poorly in terms of J, although it is optimal in terms of n^* when $n = 1$ and nothing is known about $\mu(\cdot)$ (S_3 performs better than S_2 but assumes that $\mu(\cdot)$ is known), the theoretical expected value of n^* is then $(r^* - 1)c_{r^*-1}/N \simeq 0.3743$.

Selecting one (resp. ten) X_k at random gives on the average $J = E\{X\} = \sqrt{2}/\sqrt{\pi} \simeq 0.798$ and $E\{n^*\} = 1/50$ (resp. $J = 10E\{X\} \simeq 7.98$ and $E\{n^*\} = 2$, see Rényi (1966), p. 35). During the 1000 repetitions of the experiment with $n = 10$, the 10 best items $X^*_{1,50}, \ldots, X^*_{10,50}$ were selected 116 times by S_3, 17 times by S_4 and 9 by S_1; S_2 never succeeded to select all of them. These results might seem rather poor, but it should be noticed that the probability of selecting these 10 largest values by a pure random choice is as small as $0.97\ 10^{-10}$! (see Rényi (1966), p. 35).

A statistical analysis of the results obtained for $n = 1$, based on the method of paired comparisons, shows that the strategies can be ranked by order of increasing performance as $S_2 \prec S_1 \preceq S_4 \prec S_3$ in terms of J, and as $S_4 \prec S_1 \prec S_2 \prec S_3$ in terms of n^*, with $\preceq$ indicating a non-significant difference. The situation is much different for S_2 when $n > 1$, and a statistical analysis of the results obtained for $n = 10$, again based on the method of paired comparisons, shows that the strategies can be ranked by order of increasing performance as follows, both in terms of J and n^*: $S_2 \prec S_1 \prec S_4 \prec S_3$. Taking the computing times in Table 1 into account, together with the fact that the performances of S_4 and S_1 are close, we conclude that S_1 is a most attractive strategy.

References

ALBRIGHT, S.C. and DERMAN, C. (1972): Asymptotically Optimal Policies for the Stochastic Sequential Assignment Problem. *Management Science*, 19(1), 46–51.

ALBRIGHT, S.C. (1977): A Bayesian Approach to a Generalized House Selling Problem. *Management Science*, 24(4), 432–440.

BAR-SHALOM, Y. and TSE, E. (1974): Dual Effect, Certainty Equivalence, and Separation in Stochastic Control. *IEEE Transactions on Automatic Control*, 19(5), 494–500.

BERTSEKAS, D. (1995): *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont.

DERMAN, C., LIEBERMAN, G., and ROSS, S. (1972): A Sequential Stochastic Assignment Problem. *Management Science*, 18(7), 349–355.

EMBRECHTS, P., KLÜPPELBERG, C., and MIKOSCH, T. (1997): *Modelling Extremal Events*. Springer, Berlin.

FEL'DBAUM, A. (1965): *Optimal Control Systems*. Academic Press, New York.

FELLER, W. (1966): *An Introduction to Probability Theory and Its Applications*. Wiley, New York.

FERGUSON, T. (1989): Who Solved the Secretary Problem? *Statistical Science*, 4(3), 282–296.

FREEMAN, P. (1983). The Secretary Problem and its Extensions: A Review. *International Statistical Review*, 51, 189–206.

KADISON, R. (1994): Strategies in the Secretary Problem. *Expo. Math.*, 12, 125–144.

LINDLEY, D. (1961): Dynamic Programming and Decision Theory. *Applied Statist.*, 10, 39–51.

PATCHELL, J. and JACOBS, O. (1971): Separability, Neutrality and Certainty Equivalence. *International Journal of Control*, 13(2), 337–342.

PRONZATO, L. (2001): Optimal and Asymptotically Optimal Decision Rules for Sequential Screening and Resource Allocation. *IEEE Transactions on Automatic Control*, (to appear).

RÉNYI, A. (1966): *Calcul des Probabilités*. Dunod, Paris.