Journal of the Operations Research
Society of Japan
Vol. 28, No. 3, September 1985

OPTIMAL ASSIGNMENT FOR A RANDOM SEQUENCE WITH AN UNKNOWN NUMBER OF JOBS

Tōru Nakai
Osaka Prefectural University

(Received December 10, 1982: Final February 18, 1985)

Abstract We will consider a sequential stochastic assignment problem where the number of jobs is not known beforehand. Unlike the well known sequential stochastic assignment problem, since the number of jobs is unknown, a policy of the decision-maker depends not only on the size of each job, but also on information about the number of remaining jobs. The optimal policy and the total expected reward under this policy are determined by a system of recursive equations which are obtained in the main theorem. In the last section, we will consider a case over an infinite horizon.

1. Introduction

In relation to sequential stochastic assignment problems discussed by Derman, Lieberman and Ross [2], Nakai [4] etc., we consider a sequential decision problem with an unknown number of jobs. In the sequential stochastic assignment problem, jobs arrive in sequential order, i.e., first job 1 appears, followed by job 2, etc. Here we treat a case where the number of jobs is not known in advance to the decision-maker.

We consider the following situation. There are N jobs under consideration, where N is a random variable which represents the number of remaining jobs for decisions. We assume that the probability distribution of N is given beforehand. The arrival time of each job is an independently and identically distributed random variable with a known mean. Information about the number of remaining jobs is updated in a Bayesian manner as the successive jobs are observed.

Under the above situation, a sequential stochastic assignment problem treated here is characterized by following four things. 1). The planning time period T at the last job offer, 2) The passage time t since the last job offer, i.e., the remaining time period in this situation is $T - t$ units of time. 3) Information q about the number of remaining jobs, which is improved at the last job offer. All information is summarized by a probability distribution on the set of possible numbers of jobs. Here we assume that P $\{ N \leqq M \mid q \} = 1$ for a given positive M. 4) The set of available actions $\{ p_1, \ldots, p_n \}$. Similarly to the problem considered in [2], it is assumed that $n = M$ without loss of generality. Under the above conditions, we consider the $(p_1, \ldots, p_n;\ T, t, q)$ as the state variable. Whenever a job arrives, a decision based on $(p_1, \ldots, p_n;\ T, t, q)$ is made by the decision-maker, and, therefore, we treat this problem by choosing these points of time, so as to exploit the lack of memory of the exponential distribution.

A sequential stochastic assignment problem treated here is played in the following manner. Whenever a job arrives at time t since the last job offer, i.e., the time t is an interarrival time of this job, the decision-maker updates information about the number of remaining jobs. After observing a realized value x of the random variable X associated with this job, he takes one of n available actions, where X is a size of each job and i.i.d. random variable. If the i-th action p_i is taken, an immediate reward of $p_i x$ is obtained and this action is unavailable for future decisions. Therefore, we then face a problem equivalent to one that starts in $(p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n;\ T-t, 0, \bar{q})$, where $\bar{q}$ is posterior information about the number of remaining jobs and is obtained in Section 2. Since information about the number of remaining jobs is obtained through the interarrival times of jobs, when $t = 0$, we consider that the problem is in the initial condition. Whenever the set of available actions is empty or $T = 0$, this problem stops and an immediate reward of 0 is obtained. The objective of this problem is to maximize the total expected reward.

There are several related problems. A sequential stochastic assignment problem in homogeneous Poisson arrival case is considered in Sakaguchi [5]. Sakaguchi and Tamaki concern an optimal stopping problem in a non-homogeneous Poisson arrival case in [7]. Moreover in [8], Stewart considers an optimal stopping problem in a non-homogeneous Poisson arrival case with an unknown number of options. He treats an optimal stopping problem for the relative rank. Here we consider a sequential stochastic assignment problem in a non-homogeneous Poisson arrival case with an unknown number of jobs. Sakaguchi treats a similar problem in [6].

In the following sections, we formulate the above problem by dynamic programming and state the main results in Section 3, where a system of recursive equations is obtained, and a simple example will be shown. In the final section we consider a problem over an infinite horizon.

2. Formulation of the problem

A sequential stochastic assignment problem with an unknown number of jobs in state $(p_1,\ldots,p_n;\ T,t,q)$ is formulated as follows.

There are N jobs remaining, and N is a random variable whose distribution is given by q beforehand. Regarding the number of remaining jobs, all information is summarized by a probability distribution on the set of possible numbers of jobs as $q = (q_0,q_1,\ldots,q_n)$. In this paper, we assume that $M = n$ as described in the preceding section.

Consider that N jobs are labelled $1,2,\ldots,N$. Let Z_j be an arrival time of the job labelled j, and it is assumed that $Z_1,\ldots,Z_N$ are i.i.d. exponential random variables with a known mean $1/\lambda$, i.e.,

$$\mathrm{P}\ \{\ Z_j \leqq t\ \} = 1 - \exp(-\lambda t).\quad (\ j = 1,2,\ldots,N\ )$$

Therefore the first arrival time of the job is distributed as min $\{\ Z_1,\ldots,$ $Z_n\ \}$, i.e., exponential with a mean $1/(N\lambda)$.

Let X_j, $j = 1,2,\ldots,N$, be a size of the job labelled j. The X's are i.i.d. non-negative random variables with a common c.d.f. $F(x)$ which is assumed to be known and $\mu \equiv \mathrm{E}(X_j) < \infty$.

In regard to actions p's, a value p of an action is considered as an ability of this action. This means that if the action p is chosen and assigned to any job with a realized value x, an immediate reward is given by px. That is; if $p = 1$, by using this action, the decision-maker gets the complete value x, and if $p < 1$, he gets the less value than in the case $p = 1$. It is assumed that $1 \geqq p_1 \geqq p_2 \geqq \ldots \geqq p_n \geqq 0$ for any set of available actions. Similarly to the case [2], the assumption $p \leqq 1$ is not essential. The objective of this problem is to find the optimal policy which maximizes the total expected reward.

Let $P_n(p_1,\ldots,p_n;\ T,t,q)$ be the problem in state $(p_1,\ldots,p_n;\ T,t,q)$ and the total expected reward obtainable under the optimal policy be $v_n(p_1,\ldots,p_n;\ T,t,q)$. Whenever a job arrives at time t since the last job offer, i.e., the interarrival time of this job is t, the posterior probability distribution q of N is derived from this observed interarrival time and prior information q about the number of remaining jobs, where q

is obtained by Equation (2). After observing a realized value x of the random variable X associated with this job, one of the available actions is chosen. If the i-th action p_i is chosen and assigned to this job, an immediated reward of $p_i x$ is obtained. The selected action p_i is unavailable for future decisions. We then face a problem equivalent to one that starts in $(p_1,\dots,p_{i-1},p_{i+1},\dots,p_n;\ T-t,0,\bar{q})$. These steps are repeated again and again, and this problem stops whenever the set of available actions is empty or $T = t$.

Since the arrival time of each job is i.i.d. exponential random variable with known mean $1/\lambda$, the first arrival time of the job is distributed as min $\{ Z_1,\dots,Z_n \}$, i.e., exponential with mean $1/(N\lambda)$. From the memoryless property of exponential distribution, we find that the first arrival time Y is distributed as

(1) $\mathrm{P}\{ Y \leqq t \mid N = k \} = 1 - [\exp(-\lambda t)]^k \quad (k = 1,2,\dots,n).$

By a simple application of the Bayes' theorem, (see for example DeGroot [1]), offered information $\bar{q}$ about the number of remaining jobs is given by

(2) $\bar{q}_k = c q_{k+1}(k+1)\exp(-k\lambda t)$

where $k = 0,1,\dots,n-1$, $\bar{q}_k = \mathrm{P}\{ k \text{ jobs remain} \mid \bar{q} \}$ and c is a normalizing constant to ensure that:

$$\sum_{k=0}^{n-1}\bar{q}_{k+1} = 1$$

for all t.

Since information about the number of remaining jobs is obtained through the interarrival times of jobs, here we consider no offered information q^* concerning the case where there is no job for the past t units of time since the last job offer. Analogously in above considerations, no offered information q^* about the number of remaining jobs is given by

(3) $q^*_k = d q_k \exp(-k\lambda t) \quad (k = 1,2,\dots,n)$

and $q^*_0 = dq_0$ where $q^*_k = \mathrm{P}\{ k \text{ jobs remain} \mid q^* \}$ and d is a normalizing constant to ensure that:

$$\sum_{k=0}^{n} q^*_k = 1$$

for all t. Finally we point out the fact that offered information $\bar{q}$ and no offered information q^* are functions of t.

3. Main theorem

Here we formulate this problem by dynamic programming in the following manner. In this problem, the number of jobs decreases one by one as the jobs arrive, and the rate of an arrival time is independent of time t between two successive jobs. For the problem $P_n(p_1,\ldots,p_n;\ T,t,q)$, under the conditions that $N = k$ and initial information about the number of remaining jobs is q, let $v_n^k(p_1,\ldots,p_n;\ T,t,q)$ be the total conditional expected reward as

$$(4)\quad v_n^k(p_1,\ldots,p_n;\ T,t,q) = E\ [\ v_n(p_1,\ldots,p_n;\ T,t,q)\ |\ N = k \text{ and } q\ ].$$

Therefore we have by taking expectation with respect to no offered information q^* at time t since the last job offer,

$$(5)\quad v_n(p_1,\ldots,p_n;\ T,t,q) = E\ [\ v_n^N(p_1,\ldots,p_n;\ T,t,q)\ ]$$

$$= \sum_{k=1}^{n} q_k^* v_n^k(p_1,\ldots,p_n;\ T,t,q).$$

Since information about the number of remaining jobs is updated as the successive jobs are observed, the conditional reward $v_n^k(p_1,\ldots,p_n;\ T,t,q)$ is also dependent on q.

Now we consider three cases what happen in some small time Δt when $N = k$. The first case is that, with probability $k\lambda\Delta t + o(\Delta t)$, a job arrives in Δt. When a job arrives with some observed value x, the optimal policy will be considered. The second case is that, with probability $1 - k\lambda\Delta t + o(\Delta t)$, no job arrives in Δt. The last case is that more than one job arrive in Δt, and the probability of this event is $o(\Delta t)$.

Therefore, we have, when $N = k$,

$$(6)\quad v_n^k(p_1,\ldots,p_n;\ T,t,q) = k\lambda\Delta t \int_0^\infty \max_{1\le i\le n}\{\ p_i x$$

$$+ E\ [\ v_{n-1}^{N-1}(p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_n;\ T-t-\Delta t,0,\bar{q})\ ]\ \}\ dF(x)$$

$$+ (\ 1 - k\lambda\Delta t\ )v_n^k(p_1,\ldots,p_n;\ T,t+\Delta t,q) + o(\Delta t),$$

since, whenever a job arrives at time t since the last job offer with a realized value x, a decision based on $(p_1,\ldots,p_n;\ T-t,0,\bar{q})$ is made by the decision-maker. The first term of the right hand side of Equation (6) is the first case, and the second term corresponds to the second case.

Since

$$E\ [\ v_{n-1}^{N-1}(p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_n;\ T-t,0,\bar{q})\ ]$$

$$= v_{n-1}(p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_n;\ T-t,0,\bar{q})$$

by Equation (5), rearranging terms and taking $\Delta t \to 0$, yield

$$(7)\quad \frac{\partial}{\partial t} v_n^k(p_1,\ldots,p_n;\ T,t,q) = -\ k\lambda \int_0^\infty \max_{1\le i\le n}\{\ p_i x$$

$$+ v_{n-1}(p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_n;\ T-t,0,\bar{q})\ \}\ dF(x)$$

$$+ k\lambda\ v_n^k(p_1,\ldots,p_n;\ T,t,q),$$

with the boundary condition that

$v_n^k(p_1,\ldots,p_n;\ T,t,q) = 0.$

Here we note the following thing. Since we formulate this problem by dynamic programming in Equation (6) and information about the number of remaining jobs is obtained through the interarrival times of successive jobs, we use the parameter t as an element of the state variable. On the other hand, the decision-maker can take one of available actions whenever a job arrives, and, therefore, the optimal policy is considered only at these points of time. Although information about the number of remaining jobs is also obtained through the fact that there is no job for the past t units of time since the last job offer, this informations is updated only at a point of time when a job arrives. Therefore, information q is independent of time t between two successive jobs.

The optimal policy and the total expected reward obtainable under this policy, which are determined by a system of recursive equations, are embodied in the following theorem.

Theorem 1. There exists a sequence of non-negative functions of T and t ($T,\ t \geqq 0$),

(8) $h_1^n(T,t,q) \geqq h_2^n(T,t,q) \geqq \cdots \geqq h_n^n(T,t,q) \geqq 0,$

such that the following properties are true for the problem in state $(p_1,\ldots,p_n;\ T,t,q)$.

1) Whenever a job arrives at time t since the last job offer with a realized value x, the optimal decision is as follows.

"If $h_i^{n-1}(T-t,0,\bar{q}) \leqq x < h_{i-1}^{n-1}(T-t,0,\bar{q})$ then choose the i-th action p_i and assign to it"

where $i = 1,2,\ldots,n$, $h_0^{n-1}(T-t,0,\bar{q}) = \infty$ and $h_n^{n-1}(T-t,0,\bar{q}) = 0$.

2) $h_i^n(T,t,q)$ satisfies the following system of recursive equations.

(9) $h_i^n(T,t,q) = \sum_{k=1}^{n} q_k^* \, g_i^{n,k}(T,t,q),$

where

(10) $$g_i^{n,k}(T,t,q) = k\lambda \exp(k\lambda t) \int_t^T f_i^n(T,t,q)\exp(-k\lambda t)dt$$

and

(11) $$f_i^n(T,t,q) = \int_{h_i}^{h_{i-1}} x\,dF(x) + h_{i-1}(\ 1 - F(h_{i-1})\ ) + h_i F(h_i),$$

with

$h_i = h_i^{n-1}(T-t,0,\bar{q})$

for $i = 1,2,\ldots,n$, and we define that $0\cdot\infty = 0$.

3) The value $v_n(p_1,\ldots,p_n;\ T,t,q)$ satisfies

$$v_n(p_1,\ldots,p_n;\ T,t,q) = \sum_{j=1}^{n} p_j h_j^n(T,t,q).$$

Proof: We employ the induction priciple on n. When $n = 1$, Equation (7) is

(12) $$\frac{\partial}{\partial t} v_1^1(p_1;\ T,t,q) = -\lambda \int_0^\infty p_1 x dF(x) + \lambda v_1^1(p_1;\ T,t,q).$$

The first term of the right hand side of Equation (12) is equal to $p_1\lambda\mu$, and combining the boundary condition that

$$v_1^1(p_1;\ T,T,q) = 0,$$

we have

$$v_1^1(p_1;\ T,t,q) = p_1\mu(\ 1 - \exp[-\lambda(T-t)]\ ) = p_1 g_1^{1,1}(T,t,q),$$

and

$$v_1(p_1;\ T,t,q) = E\ [\ v_1^N(p_1;\ T,t,q)\ ] = p_1 q_1^* g_1^{1,1}(T,t,q) = p_1 h_1^1(T,t,q).$$

Assume that the all parts of this theorem are true for all $m \leq n-1$. The first term of the right hand side of Equation (7) is

(13) $$-k\lambda \int_0^\infty \max_{1\leq i\leq n}\{\ \phi_i(x)\ \}\ dF(x),$$

where

(14) $$\phi_i(x) = p_i x + v_{n-1}(p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_n;\ T-t,0,\bar{q})$$
$$= \sum_{j=1}^{i-1} p_j h_j^{n-1}(T-t,0,\bar{q}) + p_i x + \sum_{j=i+1}^{n} p_j h_{j-1}^{n-1}(T-t,0,\bar{q}).$$

Equation (14) is derived from the induction assumption 3). The inequality (8) of the functions $h_j^{n-1}(T-t,0,\bar{q})$ and the well known Hardy's lemma ([2] and [3], etc) yield

$$\max_{1\leq i\leq n}\{\ \phi_i(x)\ \} = \phi_i(x) \quad \text{if} \quad h_i^{n-1}(T-t,0,\bar{q}) \leq x < h_{i-1}^{n-1}(T-t,0,\bar{q}),$$

where $i = 1,2,\ldots,$. Therefore Equation (13) is

(15) $$-k\lambda \sum_{j=1}^{n} \int_{h_j}^{h_{j-1}} \phi_j(x) dF(x)$$

where $h_j = h_j^{n-1}(T-t,0,\bar{q})$ ($j = 1,2,\ldots,n$) and $h_n^{n-1}(T-t,0,\bar{q}) = 0$. Substituting Equation (14) into Equation (15) and rearranging the terms yield

(16) $$-k\lambda \sum_{j=1}^{n} p_j f_j^n(T,t,q).$$

The solution of the differential equation (7) of this case is expressed as Equation (10), i.e.,

$$v_n^k(p_1,\ldots,p_n;\ T,t,q) = \sum_{j=1}^{n} p_j g_j^{n,N}(T,t,q).$$

Therefore, from Equation (5), we have

$$v_n(p_1,\ldots,p_n;\ T,t,q) = E\ [\ v_n^N(p_1,\ldots,p_n;\ T,t,q)\ ]$$
$$= \sum_{j=1}^{n} p_j h_j^n(T,t,q).$$

On the other hand, the inductive hypothesis,

$$h_1^{n-1}(T-t,0,\bar{q}) \geqq h_2^{n-1}(T-t,0,\bar{q}) \geqq \dots \geqq h_{n-1}^{n-1}(T-t,0,\bar{q}),$$

yields

(17) $f_1^n(T,t,q) \geqq f_2^n(T,t,q) \geqq \dots \geqq f_n^n(T,t,q)$.

This inequality is obtained from a simple calculation. Inequality (17) and the fact that $g_i^{n,k}(T,t,q) = y(t)$ is the solution of the differential equation

$$\frac{\partial}{\partial t} y(t) - k\lambda y(t) = - k\lambda f_i^n(T,t,q),$$

yield $g_i^{n,k}(T,t,q) \geqq g_{i-1}^{n,k}(T,t,q) \geqq 0$. (i, $k = 1,2,\dots,n$) Therefore we have

$$h_i^n(T,t,q) \geqq h_{i-1}^n(T,t,q). \qquad (i = 1,2,\dots,n)$$

This completes the proof of this theorem

Concerning Theorem 1, here we note that $h_i^n(T,t,q)$ and $g_i^{n,k}(T,t,q)$ (i, $k = 1,2,\dots,n$) are increasing in t and decreasing in T. These things are derived from Equations (9) and (10).

The result of Theorem 1 is, in form, similar to one of Sakaguchi [5] because of the similar situation that the number of jobs is unknown. The problem in [5] is a Poisson arrival case and the problem of this paper is a generalization to a non-homogeneous Poisson arrival case. The difference comes from the fact that this problem has only a limited number of jobs, contrary to the problem with an unlimited number of jobs in [5]. Concerning the problem in [5], the arrival rate of jobs is the same under any situations. However, the problem considered in this paper concerns the case that the value of the problem depends not only on information q for the number of remaining jobs, but also on the interarrival time of jobs. Since the number of jobs is not known in advance, the decision-maker obtains information for the number of remaining jobs from interarrival times of jobs. In this problem, the precise number of jobs becomes known later to the decision-maker. The difficulty of this problem arises from these facts. Concerning the similar situation, an optimal stopping problem for the relative rank is considered in Stewart [8].

Concerning the total expected reward $v_n(p_1,\dots,p_n;\ T,t,q)$, here we consider the condition that a job arrives at time t since the last job offer. Under this condition, the conditional value of this problem is

$$\sum_{j=1}^{n} p_j f_j^n(T,t,q),$$

which is obtained in Equation (16) of Theorem 1. Moreover, we note the fact that $f_j^n(T,t,q) = f_j^n(T-t,0,\bar{q})$ by Equation (11). On the other hand, the value $v_n(p_1,\dots,p_n;\ T,t,q)$ represents the total expected reward under the optimal policy of the problem in $(p_1,\ldots,p_n;\ T,t,q)$, i.e.,

$$\sum_{k=1}^{n} q_k^* k\lambda e^{k\lambda t} \int_t^T \left(\sum_{j=1}^{n} p_j f_j^n(T,t,q) \right) e^{-k\lambda t} dt,$$

where q^* is no offered information about the number of remaining jobs at the point of time t since the last job offer when no job arrives for the past t units of time and at time t. Moreover, when $t = 0$, the value $v_n(p_1,\ldots,p_n;\ T,0,q)$ is

$$\sum_{k=1}^{n} q_k k\lambda \int_0^T \left(\sum_{j=1}^{n} p_j f_j^n(T,t,q) \right) e^{-k\lambda t} dt.$$

In relation to the problem considered here, next we observe a simple example in the following manner.

Example 1. We assume that the random variable X which represents a size of each job, is uniform on [0,1], i.e.,

$$F(x) = x \quad (0 \leqq x \leqq 1).$$

First we consider the case with $n = 1$. Since $f_1^1(T,t,q) = E[\,X\,] = 1/2$,

$$h_1^1(T,t,q) = q_1^* g_1^{1,1}(T,t,q)$$

and

$$g_1^{1,1}(T,t,q) = \frac{\lambda}{2} e^{\lambda t} \int_t^T e^{-\lambda t} dt = (1 - e^{-\lambda(T-t)})/2.$$

Here we note that $h_1^1(T,t,q)$ is decreasing in t, increasing in T and $h_1^1(T,T,q) = 0$.

Next we consider the case with $n = 2$. In this case, Theorem 1 yields

$$(18)\quad h_i^2(T,t,q) = \sum_{k=1}^{2} q_k^* g_i^{2,k}(T,t,q), \quad (i = 1,2)$$

where

$$(19)\quad g_i^{2,k}(T,t,q) = k\lambda\, e^{k\lambda t} \int_t^T f_i^2(T,t,q) e^{-k\lambda t} dt \quad (k = 1,2)$$

and

$$f_1^2(T,t,q) = \int_0^{h_1^1} h_1^1 dF(x) + \int_{h_1^1}^{\infty} x dF(x) = h_1^1 + T_F(h_1^1),$$

$$f_2^2(T,t,q) = \int_0^{h_1^1} x dF(x) + \int_{h_1^1}^{\infty} h_1^1 dF(x) = 1/2 - T_F(h_1^1),$$

where $h_1^1 = h_1^1(T-t,0,\bar{q})$ and $T_F(a) = \int_a^{\infty} (x-a) dF(x)$. The function $T_F(a)$ is a well known function in the decision analysis (see DeGroot [1]), and in this case

$$T_F(a) = \frac{1}{2} (1 - a)^2.$$

Since $\bar{q} = (cq_1,\ 2cq_2 e^{-\lambda t})$ where $c = (q_1 + 2q_2 e^{-\lambda t})^{-1}$, we have

$$h_1^1 = h_1^1(T-t,0,\bar{q}) = 2cq_2 e^{-\lambda t}(1 - e^{-\lambda(T-t)})/2 = cq_2(e^{-\lambda t} - e^{-\lambda T}).$$

Therefore, whenever a job arrives with a realized value x at time t since the last job offer, the optimal decision is to

" assign p_1 if $h_1^1 \leqq x$, and assign p_2 if $x < h_1^1$ "

when the problem is in $(p_1,p_2;\ T,t,q)$.

In order to obtain the value $v_2(p_1,p_2;\ T,t,q)$ for the case with $n = 2$, we need to calculate the value of

$$\int_t^T T_F(h_1^1(T-t,0,\bar{q}))e^{-i\lambda t}dt \qquad (\ i = 1,2\ )$$

in the following manner.

$$\int T_F(h_1^1(T-t,0,\bar{q}))e^{-\lambda t}dt$$
$$= -\{\ e^{-\lambda t} + (\alpha(T)/q_2)\log(\alpha(t)) - \alpha(T)^2/2q_2\alpha(t)\ \}/8\lambda,$$

and

$$\int T_F(h_1^1(T-t,0,\bar{q}))e^{-2\lambda t}dt$$
$$= -\{\ q_2e^{-2\lambda t} + 2\alpha(T)e^{-\lambda t} + (\ \alpha(T)(\alpha(T)-2q_1)/2q_2\ )\log(\alpha(t))$$
$$+ q_1\alpha(T)^2/2q_2\alpha(t)\ \}/16q_2\lambda,$$

where

$$\alpha(t) = q_1 + 2q_2e^{-\lambda t}.$$

In order to obtain the optimal policy for the case with $n = 3$, we consider a special case with $t = 0$. Therefore from Equation (9), we get the values in the following manner, i.e.,

$$h_i^2(T,0,q) = \sum_{k=1}^2 q_k g_i^{2,k}(T,0,q)$$
$$= \sum_{k=1}^2 k\lambda q_k \int_0^T f_i^2(T-t,0,\bar{q})e^{-k\lambda t}dt$$
$$= \int_0^T f_i^2(T-t,0,\bar{q}) \sum_{k=1}^2 k\lambda q_k e^{-k\lambda t}dt. \qquad (\ i = 1,2\ )$$

First we get from the above calculations,

$$\xi(T,q) = \int_0^T T_F(h_1^1(T-t,0,\bar{q}))(\lambda q_1e^{-\lambda t} + 2\lambda q_2e^{-2\lambda t})dt$$
$$= \int_0^T \frac{1}{2}\ (1 - q_2(e^{-\lambda t} - e^{-\lambda T})/(q_1+2q_2e^{-\lambda t}))^2\lambda e^{-\lambda t}(q_1+2q_2e^{-\lambda t})dt$$
$$= a_2e^{-2\lambda T} + a_1e^{-\lambda T} + a_0 + a_{-1}\log(\alpha(0)/\alpha(T)),$$

where

$$a_2 = -5q_2/8,$$
$$a_1 = (4q_2-3q_1)/8,$$
$$a_0 = (q_2+3q_1)/8,$$

$a_{-1} = \alpha(T)^2/16q_2$.

Therefore we have

$$h_2^2(T,0,q) = \frac{1}{2}\int_0^T (\lambda q_1 e^{-\lambda t} + 2\lambda q_2 e^{-2\lambda t})dt - \xi(T,q)$$
$$= \frac{1}{2}(\ q_1 + q_2 - q_1 e^{-\lambda T} - q_2 e^{-2\lambda T}\ ) - \xi(T,q),$$
$$h_1^2(T,0,q) = \int_0^T q_2(e^{-\lambda t} - e^{-\lambda T})\lambda e^{-\lambda t}dt + \xi(T,q)$$
$$= q_2\ (\ 1/2 + e^{-2\lambda T}/2 - e^{-\lambda T}\ ) + \xi(T,q),$$

since $cq_2(e^{-\lambda t} - e^{-\lambda T})(q_1\lambda e^{-\lambda t} + 2q_2\lambda e^{-2\lambda t}) = q_2(e^{-\lambda t} - e^{-\lambda T})\lambda e^{-\lambda t}$.

The value $v_2(p_1,p_2;\ T,0,q)$ of the problem $P_2(p_1,p_2;\ T,0,q)$ is given by

$$v_2(p_1,p_2;\ T,0,q) = p_1h_1^2(T,0,q) + p_2h_2^2(T,0,q).$$

Values for other cases with $n = 2$, are obtained similarly.

On the other hand, when the problem is in $(p_1,p_2,p_3;\ T,t,q)$, if a job arrives with a realized value x at time t since the last job offer, the optimal decision is to;

$$\text{assign}\ \begin{Bmatrix} p_1 \\ p_2 \\ p_3 \end{Bmatrix}\ \text{if}\ \begin{cases} x \geqq h_1^2(T-t,0,\bar{q}) \\ h_1^2(T-t,0,\bar{q}) > x \geqq h_2^2(T-t,0,\bar{q}) \\ h_2^2(T-t,0,\bar{q}) > x \geqq 0. \end{cases}$$

Treating for $n \geqq 3$, it is extremely complicated, and we omit it here. Moreover if we consider a special case where $q_1 = q_2 = 1/2$, this example is equivalent to an example treated in [6].

4. Infinite horizon case

In this section we consider a problem in the infinite horizon case, i.e., the period T of the problem is not restricted. Concerning this case, we assume a discount factor $\beta \geqq 0$. In this section we use, for the state, the notation $(p_1,\ldots,p_n;\ t,q)$ instead of $(p_1,\ldots,p_n;\ T,t,q)$ used in the preceding sections. From an argument similar to one used in the last section, we have the following differential equation

(20) $$v_n(p_1,\ldots,p_n;\ t,q) = E\ [\ v_n^N(p_1,\ldots,p_n;\ t,q)\ ],$$

(21) $$\frac{\partial}{\partial t} v_n^k(p_1,\ldots,p_n;\ t,q) = -\ k\lambda\int_0^\infty \max_{1\leq i\leq n}\ \{\ \psi_i(x)\ \}\ dF(x) + (\ k\lambda + \beta\ )v_n^k(p_1,\ldots,p_n;\ t,q),$$

where $\psi_i(x) = p_i x + v_{n-1}(p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_n;\ 0,\bar{q})$.
Therefore we have the following theorem, and the proof of this theorem is obtained through a method similar to one used in Theorem 1. If $\beta = 0$, the differential equation (21) is equivalent to Equation (7).

Theorem 2. For any t and q, there exists a sequence of non-negative functions of t ($\geqq 0$),

(22) $\bar{h}_1^n(t,q) \geqq \bar{h}_2^n(t,q) \geqq \ldots \geqq \bar{h}_n^n(t,q) \geqq 0$,

such that the following three facts are true for the problem $P_n(p_1,\ldots,p_n;\ t,q)$.

1) When the decision-maker observes a job with a realized value x at time t since the last job offer, the optimal decision is to

" choose the i-th action p_i if $\bar{h}_i^{n-1}(0,\bar{q}) \leqq x < \bar{h}_{i-1}^{n-1}(0,\bar{q})$ "

where $i = 1,2,\ldots,n$, $\bar{h}_n^{n-1}(0,\bar{q}) = \infty$ and $\bar{h}_n^{n-1}(0,\bar{q}) = 0$.

2) $\bar{h}_i^n(t,q)$ satisfies

$$\bar{h}_i^n(t,q) = \sum_{k=1}^{n} q_k \bar{g}_i^{n,k}(t,q),$$

(23) $$\bar{g}_i^{n,k}(t,q) = k\lambda\exp[(k\lambda+\beta)t]\int_t^\infty \bar{f}_i^n(t,q)\exp[-(k\lambda+\beta)t]\mathrm{d}t,$$

$$\bar{f}_i^n(t,q) = \int_{h_i}^{h_{i-1}} x\mathrm{d}F(x) + h_{i-1}(1-F(h_{i-1})) + h_i F(h_i),$$

where $h_i = \bar{h}_i^{n-1}(0,\bar{q})$ ($i = 1,2,\ldots,n$), and we define that $0\cdot\infty = 0$.

3) We have

$$v_n(p_1,\ldots,p_n;\ t,q) = \sum_{j=1}^{n} p_j \bar{h}_j^n(t,q).$$

Concerning the infinite horizon case, here we point out the following relation to the problem where the number of jobs is known to the decision-maker previously. In the problem [7], if we consider the infinite horizon case, the result of this work is the same to one in the optimal stopping problem where the number of jobs is a known and fixed constant. This comes from the fact that every job will certainly arrive and the decision-maker is able to set his hope on remaining jobs, i.e., this problem is equivalent to one considered in Derman, Lieberman and Ross [2]. However, in our problem considered here, the decision-maker only knows the prior probability distribution about the number of remaining jobs, and does not know the precise number of jobs. The decision-maker always guesses the number of remaining jobs, i.e., information about the number of remaining jobs is updated as the interarrival times of successive jobs are observed. The difficulty of this problem arises from this fact. Moreover, since the number of jobs is less than or equal to the number of available actions by our assumption, there might be several actions which are not assigned to any job.

Finally we shall reconsider Example 1 for the problem of this section.

Example 2. Under the same conditions of Example 1, we consider the case with $n = 1$, then we have

$$\bar{g}_1^{1,1}(t,q) = (\ \lambda/2(\lambda+\beta)\ )e^{-\beta t}$$

and

$$\bar{h}_1^1(t,q) = q_1^*\bar{g}_1^{1,1}(t,q) = q_1^*(\lambda/2(\lambda+\beta))e^{-\beta t}.$$

Next we consider the case with $n = 2$. Similarly to Example 1, we have

$$\bar{h}_i^2(t,q) = \sum_{k=1}^{2} q_k^*\bar{g}_i^{2,k}(t,q), \qquad (\ i = 1,2\ )$$

where

$$\bar{g}_i^{2,k}(t,q) = k\lambda e^{(k\lambda+\beta)t}\int_t^{\infty} \bar{f}_i^2(t,q)e^{-(k\lambda+\beta)t}dt \qquad (\ k = 1,2\ )$$

and

$$\bar{f}_1^2(t,q) = \bar{h}_1^1 + T_F(\bar{h}_1^1), \qquad \bar{f}_2^2(t,q) = 1/2 - T_F(\bar{h}_1^1),$$

$$\bar{h}_1^1 = \bar{h}_1^1(0,\bar{q}) = cq_2e^{-\lambda t}\times(\lambda/(\lambda+\beta)) \qquad (\ c = (\ q_1 + 2q_2e^{-\lambda t}\ )^{-1}\ ).$$

Therefore, whenever a job arrives with a realized value x at time t since the last job offer, the optimal decision is to

" assign p_1 if $\bar{h}_1^1 \leqq x$, and assign p_2 if $\bar{h}_1^1 > x$ "

when the problem is in $(p_1,p_2;\ t,q)$.

First we consider the case with $\beta = 0$. Similarly to Example 1, we get the values of $\bar{h}_i^2(0,q)$'s in the following manner. We have

$$\xi(q) = \int_0^{\infty} T_F(\bar{h}_1^1(0,\bar{q}))(\lambda q_1e^{-\lambda t} + 2\lambda q_2e^{-2\lambda t})dt$$

$$= (q_2+3q_1)/8 + ((q_1)^2/(16q_2))\ \log((q_1+2q_2)/q_1).$$

Therefore we have

$$\bar{h}_1^2(0,q) = q_2/2 + \xi(q)$$

and $\bar{h}_2^2(0,q) = (q_1 + q_2)/2 - \xi(q)$.

Especially, when $q_1 = 1$ and $q_2 = 0$, $\bar{h}_1^2(0,q) = 1/2$ and $\bar{h}_2^2(0,q) = 0$; when $q_1 = 0$ and $q_2 = 1$, $\bar{h}_1^2(0,q) = 5/8$ and $\bar{h}_2^2(0,q) = 3/8$; and when $q_1 = q_2 = 1/2$, $\bar{h}_1^2(0,q) = 1/2 + (\ \log 3\ )/32$ and $\bar{h}_2^2(0,q) = 1/4 - (\ \log 3\ )/32$. Hence, from Theorem 2, the value $v_2(p_1,p_2;\ 0,q)$ is given by

$$v_2(p_1,p_2;\ 0,q) = p_1\bar{h}_1^2(0,q) + p_2\bar{h}_2^2(0,q)$$

$$= p_2(q_1/2) + (p_1+p_2)(q_2/2) + (p_1-p_2)\xi(q).$$

Next we consider the case with $\beta = \lambda > 0$. Similarly to the above case, we get the values $\bar{h}_i^2(0,q)$'s as follows. Since

$$\bar{h}_1^1(0,\bar{q}) = cq_2 e^{-\lambda t}/2,$$

we have

$$\eta(q) = \int_0^\infty T_F(\bar{h}_1^1(0,\bar{q})(\lambda q_1 e^{-\lambda t} + 2\lambda q_2 e^{-2\lambda t})dt$$

$$= (12(q_2)^2+15q_1q_2+(q_1)^2)/(64q_2) - ((q_1)^3/(128(q_2)^2))\log((q_1+2q_2)/q_1)$$

Therefore we have

$$\bar{h}_1^2(0,q) = q_2/6 + \eta(q),$$

and $\bar{h}_2^2(0,q) = (3q_1+4q_2)/12 - \eta(q)$.

By a simple calculation, we get the inequality

$$\bar{h}_1^2(0,q) \geqq \bar{h}_2^2(0,q)$$

for any $q = (q_0,q_1,q_2)$. Especially, when $q_1 = 1$ and $q_2 = 0$, $\bar{h}_1^2(0,q) = 1/4$ and $\bar{h}_2^2(0,q) = 0$; when $q_1 = 0$ and $q_2 = 1$, $\bar{h}_1^2(0,q) = 17/48$ and $\bar{h}_2^2(0,q) = 7/48$; and when $q_1 = q_2 = 1/2$, $\bar{h}_1^2(0,q) = 29/96 - (\log 3)/256$ and $\bar{h}_2^2(0,q) = 7/96 + (\log 3)/256$. Here we treat a problem with a discount factor, and $\bar{h}_1^2(0,q) + \bar{h}_2^2(0,q) = q_1/4 + q_2/2$ for any q. We have

$$v_2(p_1,p_2;\ 0,q) = p_1\bar{h}_1^2(0,q) + p_2\bar{h}_2^2(0,q).$$

Acknowledgement

The author wishes to thank Professor M. Sakaguchi of Osaka University for his encouragement and guidance, and the referees for their helpful comments.

References

[1] DeGroot, M. H.: *Optimal Statistical Decisions*. McGraw-Hill, New York, 1970.

[2] Derman, C., Lieberman, G. J. and Ross, S. M.: A Sequential Stochastic Assignment Problem. *Management Science*, vol. 18 (1972), 349-355.

[3] Hardy, G. H., Littlewood, J. E. and Polya, G.: *Inequalities*, Cambridge University Press, England, 1934.

[4] Nakai, T.: Optimal Assignment for a Random Sequence with an Unknown Parameter. *Journal of Information & Optimization Science*, vol. 1 (1980),

214-228.

[5] Sakaguchi, M.: A Sequential Assignment Problem for Randomly Arriving Jobs. *Reports of Statistical Applications Research, JUSE,* vol. 19 (1972), 99-109.

[6] Sakaguchi, M.: A Sequential Stochastic Assignment Problem with an Unknown Number of Jobs. *Mathematica Japonica,* vol. 29 (1984), 141-152.

[7] Sakaguchi, M. and Tamaki, M.: Optimal Stopping Problems Associated with a Non-homogeneous Markov Process. *Mathematica Japonica,* vol. 25 (1980), 681-696.

[8] Stewart, T. J.: The Secretary Problem with an Unknown Number of Options. *Operations Research,* vol. 29 (1980), 130-145.

Tōru NAKAI
Department of Mathematics,
College of Integrated Arts and Sciences,
Osaka Prefectural University,
Sakai 591,
Osaka, Japan.