Journal of the Operations Research
Society of Japan
Vol. 30, No. 1, March 1987

VARIANCE CONSTRAINED MARKOV DECISION PROCESS

Hajime Kawai
University of Osaka Prefecture

Naoki Katoh
Kobe University of Commerce

(Received September 11, 1985; Revised August 23, 1986)

Abstract The problem considered for a Markov decision process is to find an optimal randomized policy that maximizes the expected reward in a transition in the steady state among the policies which secure the variance not larger than a specified value. A solution method by a parametric Markov decision process is developed. An optimal policy is shown to be a mixture of at most two pure policies. As an application, a sequential replacement problem of a Markovian deterioration system is discussed.

1. Introduction

The optimality criteria most commonly used in a Markov decision process (MDP) are the expected total discounted reward and the expected average reward. In many kinds of applications, a decision maker who is a risk averser may not be satisfied with just maximizing the expected reward. He is concerned with the variance, which is one of the most important criteria for risk. For the average reward case, Mandl[5] discusses the variance for the policies which maximize the average reward. For the discounting case, Sobel [6] presents the formulae for the variance of the present value and discusses the mean-variance trade off problem. Benito[1] developes a calculation method of the variance of the immediate reward in a transition in the steady state.

In this paper, we consider an optimization problem for the reward in the steady state of a MDP with no discounting. The problem is to find an optimal randomized policy that maximizes the expected reward in a transition among the policies which secure the variance not larger than a specified value.

The problem is first formulated by a nonlinear programming with the feasible region which is not a convex set. It makes the programming be troublesome to solve directly by considering the corresponding Lagrangian problem. To overcome this point, we next introduce a parametric MDP which enables us to develope a procedure to find an optimal policy of our variance constrained problem. Finally, a simple sequential replacement problem of a Markovian deterioration system is discussed as an application.

2. Markov decision process with randomized policies

We consider a discrete time Markov process with finite set of states $I=\{0,1,\ldots,m\}$. For each state i, we have a finite set of actions $K_i=\{1,2,\ldots,k_i\}$. When action k is chosen at state i, the process moves to state j at the next time with probability p_{ij}^k. When action k is selected at state i and the state at the next time is j, the immediate reward r_{ij}^k is generated. In our problem, we can assume that each r_{ij}^k is nonnegative for all i,j and k without loss of generality. In this paper, we restrict ourselves only to randomized stationary policies and assume that each nonrandomized (pure) stationary policy determines only one ergodic class. For such a MDP, we consider the following one period problem in the steady state. We let mean and variance imply the expected value and the variance of the immediate reward in a transition in the steady state, respectively. Then our problem considered here is to find a policy which maximizes the mean among the randomized stationary policies which give the variance not larger than the value V. In this section, we present the formulae for the mean and variance when a policy is fixed.

Let choose action k at state i at each time with probability d_i^k, then it is cleat that

$$(2.1) \qquad \sum_{k\in K_i} d_i^k = 1, \quad d_i^k \geqq 0, \quad k \in K_i,\ i \in I.$$

Let π_i denote the probability that the process is in state i in the steady state. Then they must obey the following equation:

$$\pi_j - \sum_{i\in I} \sum_{k\in K_i} \pi_i d_i^k p_{ij}^k = 0,$$

(2.2) $\sum_{i \in I} \pi_i = 1,$

$\pi_i \geqq 0, \quad i \in I.$

We let s_t and a_t denote the state and the action selected at time t, then the mean is given by

(2.3) $\lim_{t \to \infty} E[\, r^{a_t}_{s_t, s_{t+1}} \,] = \sum_{i \in I} \sum_{k \in K_i} \pi_i d_i^k r_i^k,$

and the variance is given by

(2.4) $\lim_{t \to \infty} \mathrm{Var}[\, r^{a_t}_{s_t, s_{t+1}} \,] = \sum_{i \in I} \sum_{k \in K_i} \pi_i d_i^k R_i^k - (\sum_{i \in I} \sum_{k \in K_i} \pi_i d_i^k r_i^k)^2,$

where

$$r_i^k = \sum_{j \in I} r_{ij}^k p_{ij}^k,$$

$$R_i^k = \sum_{j \in I} (r_{ij}^k)^2 p_{ij}^k.$$

3. Formulation by nonlinear programming

In the sequel, we let $\sum_s$ denote the sum over all the values which s can take.

Setting

(3.1) $x_j^k = \pi_j d_j^k,$

in equations (2.2)-(2.4), our problem is formulated as the following nonlinear programming:

(3.2) maximize $\sum_{jk}\sum r_j^k x_j^k,$

subject to

$$\sum_k x_j^k - \sum_{ik}\sum x_i^k p_{ij}^k = 0, \quad j=1,2,\ldots,m,$$

$$\sum_{jk}\sum x_j^k = 1, \tag{3.3}$$

$$x_j^k \geqq 0, \quad k \in K_j, \; j \in I,$$

$$\sum_{jk}\sum R_j^k x_j^k - (\sum_{jk}\sum r_j^k x_j^k)^2 \leqq V. \tag{3.4}$$

Letting $\bar{x}_j^k$ be an optimal solution if exists, optimal probabilities $\bar{d}_j^k$ are obtained from

$$\bar{x}_j^k = (\sum_k \bar{x}_j^k)\bar{d}_j^k. \tag{3.5}$$

We let P_V denote this problem.

Here, for the latter discussion, we briefly touch upon the usual minimization problem of the expected average cost. We let c_i^k be the expected immediate cost when action k is selected at state i. This problem is formulated as the following linear programming (Derman[2]).

$$\text{minimize} \quad \sum_{jk}\sum c_j^k x_j^k,$$

subject to the constraints of equation (3.3).

We let $P(c_i^k,\ i\in I,\ k\in K_i)$ denote this problem. When there is no confusion, it is abbreviated to $P(c_i^k)$. It is known that considering the dual problem of this linear programming, a policy which satisfies the following functional equation with respect to g, $v_0, v_1, \ldots, v_m$ is shown to be optimal.

$$\begin{aligned} g + v_i &= \min_{k\in K_i} \{ c_i^k + \sum_j p_{ij}^k v_j \}, \quad i \in I, \\ v_0 &= 0, \end{aligned} \tag{3.6}$$

where g and v_i are called as gain and relative value, respectively. An optimal policy is obtained by usual technique of policy improvement method (Howard[3]) or successive approximation method (White[7]). It should be noted that in the MDP only with the mean criterion, there exists a pure policy which is optimal even if the range of policy is extended to randomized policy.

4. Parametric Markov decision process

In the nonlinear programming formulation in the preceding section, it is easily seen that the left hand side of inequality (3.4) is cancave in $(x_0^1, \ldots, x_m^k)$, that is, the feasible region is not a convex set. Hence, it is usually troublesome to derive an optimal policy by solving the nonlinear programming directly. In order to avoid this difficulty, we propose a parametric MDP which enables us to determine an optimal policy without appealing to nonlinear programming formulation.

We treat the problem $P(b_i^k+\theta a_i^k)$, where θ is nonnegative real number. For this problem, we investigate the behavior of an optimal policy with respect to the parameter θ.

Letting a policy B denote an optimal policy of the problem $P(b_i^k)$, then if we use superscript B to indicate the quantities associated with the policy, policy B obeys the following equation (see equation (3.6)):

$$(4.1)\quad \begin{aligned} &g^B + v_i^B = b_i^B + \sum_j p_{ij}^B v_j^B \leq b_i^k + \sum_j p_{ij}^k v_j^B, \quad k \in K_i, \quad i \in I \\ &v_0^B = 0. \end{aligned}$$

We let H_i denote the set of optimal actions at state i, that is

$$(4.2)\quad H_i = \{k \mid k \in K_i, \quad b_i^k + \sum_j p_{ij}^k v_j^B = b_i^B + \sum_j p_{ij}^B v_j^B\}, \quad i \in I.$$

We let A denote an optimal policy of the problem $P(a_i^k,\ i\in I,\ k\in H_i)$. That is, policy A satisfies the following equation:

$$(4.3)\quad \begin{aligned} &h^A + w_i^A = a_i^A + \sum_j p_{ij}^A w_j^A \leq a_i^k + \sum_j p_{ij}^k w_j^A, \quad k \in H_i, \quad i \in I, \\ &w_0^A = 0. \end{aligned}$$

Here it should be noted that

$$(4.4)\quad g^A + v_i^A = b_i^A + \sum_j p_{ij}^A v_j^A \leq b_i^k + \sum_j p_{ij}^k v_j^A, \quad k \in K_i, \quad i \in I,$$

$$(4.5)\quad g^B = g^A, \quad v_i^B = v_i^A, \quad i \in I.$$

since policy A is also optimal for $P(b_i^k)$.

We let G^A and V_i^A denote the gain and relative values for the MDP with the expected immediate cost $b_i^k+\theta a_i^k$ which is operated under policy A, respectively. Then it holds that

$$(4.6)\quad \begin{aligned} G^A + V_i^A &= \theta a_i^A + b_i^A + \sum_j p_{ij}^A V_j^A, \quad i \in I, \\ V_0^A &= 0. \end{aligned}$$

From equations (4.1),(4.3)-(4.6), we have that

$$(4.7)\quad \begin{aligned} G^A &= \theta h^A + g^A, \\ V_i^A &= \theta w_i^A + v_i^A, \end{aligned}$$

since equation (4.6) has a unique solution.

We define α_i^k and β_i^k as follows:

$$(4.8)\quad \begin{aligned} \alpha_i^k &= a_i^A + \sum_j p_{ij}^A w_j^A - (a_i^k + \sum_j p_{ij}^k w_j^A), \\ \beta_i^k &= b_i^A + \sum_j p_{ij}^A v_j^A - (b_i^k + \sum_j p_{ij}^k v_j^A). \end{aligned}$$

From equations (4.1) and (4.3), we have that

$$(4.9)\quad \begin{aligned} \beta_i^k &\begin{cases} = 0 & \text{for } k \in H_i \\ < 0 & \text{for } k \in K_i - H_i, \end{cases} \\ \alpha_i^k &\leqq 0 \quad \text{for } k \in H_i \end{aligned}$$

We defien

$$(4.10)\quad J = \{ (i,k) \mid \alpha_i^k > 0,\ i \in I,\ k \in K_i \},$$

and

$$(4.11)\quad \bar{\theta} = \begin{cases} \min\{ -\beta_i^k/\alpha_i^k \mid (i,k) \in J\}, & \text{if } j \neq \phi, \\ \infty & , \text{ otherwise.} \end{cases}$$

From equations (4.9)-(4.11), it is easily seen that $\bar{\theta}$ is positive. From equations (4.6)-(4.11), it holds that for any $\theta \in [0,\bar{\theta}]$,

$$(4.12)\quad G^A + V_i^A \leqq \theta a_i^k + b_i^k + \sum_j p_{ij}^k V_j^A, \quad k \in K_i, \quad i \in I,$$

which implies that policy A is optimal for any $\theta \in [0,\bar{\theta}]$.

Letting $A(\theta)$ denote an optimal policy of the problem $P(b_i^k+\theta a_i^k)$, then the above discussion implies that there exist $\theta_0,\theta_1,\ldots,\theta_N$ such that

$$0 = \theta_0 < \theta_1 < \ldots < \theta_n < \ldots < \theta_N < \infty,$$

$$(4.13)\quad A(\theta) = \begin{cases} A_n & \text{for } \theta \in [\theta_n,\theta_{n+1}], \quad n=0,1,\ldots,N-1, \\ A_N & \text{for } \theta \in [\theta_N,\infty), \end{cases}$$

and that θ_n and A_n are obtained by the following procedure:

Step 0: Set $n=0$ and $\theta_n=0$.
Step 1: Determine a set H_i of optimal actions for the problem $P(b_i^k+\theta a_i^k)$.
Step 2: Determine an optimal policy A_n for the problem $P(a_i^k,\ i\in I,\ k\in H_i)$.
Step 3: Set $A=A_n$ in equations (4.3) and (4.4) and calculate v_i^A and w_i^A.
Step 4: Calculate $\bar{\theta}$ by equations (4.8)-(4.11).
Step 5: If $\bar{\theta}=\infty$, then set $N=n$ and stop. Otherwise, calculate $\theta_{n+1}=\theta_n+\bar{\theta}$, set $n=n+1$ and go back to step 1.

It should be noted that for $n\geqq 1$, we need not solve the problem $P(b_i^k+\theta_n a_i^k)$ since policy A_{n-1} is an optimal policy for this problem. Hence, step 2 starts from setting $B=A_{n-1}$ in equation (4.1).

5. Variance constrained MDP

Considering the mean minimization problem for the MDP with the expected immediate cost $\theta R_i^k - r_i^k$, policies $A_0,A_1,\ldots,A_N$ are determined in the same fashion as equation (4.13). The gain of the MDP for policy A_n is given by

$$(5.1)\quad \sum_i (\theta R_i^{An} - r_i^{An})\pi_i^{An} = \theta G_n - g_n,$$

where

$$(5.2)\quad G_n = \sum_i R_i^{An}\pi_i^{An},$$

$$g_n = \sum_i r_i^{An} \pi_i^{An}.$$

It should be noted that G_n and g_n are the square mean and the mean for policy A_n, respectively. Hence, the variance V_n for policy A_n is given by $G_n - g_n^2$.

In this section, we develope a solution method for the variance constrained problem P_V by using A_n, G_n and g_n for the parametric MDP $P(\theta R_i^k - r_i^k)$. We let G and g denote the square mean and the mean for an arbitrary fixed randomized stationary policy, which are obtained from equations (2.2)-(2.4). Then by the optimality of policy A_n for the problem $P(\theta R_i^k - r_i^k)$, where $\theta_n \leqq \theta \leqq \theta_{n+1}$, $n \leqq N-1$, we have that

$$\begin{aligned} &\theta_n G_n - g_n \leqq \theta_n G - g, \\ (5.3) \quad &\theta_{n+1} G_{n+1} - g_{n+1} = \theta_{n+1} G_n - g_n \leqq \theta_{n+1} G - g. \end{aligned}$$

For the behavior of G_n and g_n with respect to n, we have the following lemma.

Lemma 1. G_n and g_n are nonincreasing in n.

Proof: Using equation (5.3), the proof is easily done.

We let $\bar{G}$ and $\bar{g}$ denote the square mean and the mean for an optimal policy of the problem P_V if exists.

Lemma 2. If $V_0 \leqq V$, then policy A_0 is optimal.

Proof: From equation (5.3), we have that $-g_0 \leqq -\bar{g}$.

Here, we treat the case where $V_0 > V$, that is, policy A_0 is not a feasible one.

Lemma 3. It holds that $G_N \leqq \bar{G} < G_0$.

Proof: It holds that $\bar{g} \leqq g_0$, $\bar{G} - \bar{g}^2 \leqq V < G_0 - g_0^2$, which implies that $0 \leqq g_0^2 - \bar{g}^2 < G_0 - \bar{G}$. From equation (5.3), we have that for any $\theta \geqq \theta_N$, $\theta G_N - g_N \leqq \theta \bar{G} - \bar{g}$, which implies that $G_N \leqq G$.

Lemma 4. If $G_{n+1} \leqq \bar{G} \leqq G_n$, then (i) $V_n \leqq V$ or (ii) $V_n > V$ and $V_{n+1} \leqq V$.

Proof: We assume that $V_n > V$ and $V_{n+1} > V$. From equation (5.3), we have that

$$\bar{G} - \bar{g}^2 \geqq f(\bar{G}) \equiv \bar{G} - \{\theta_{n+1}(\bar{G} - G_{n+1}) + g_{n+1}\}^2,$$

$$f(G_{n+1}) = G_{n+1} - g_{n+1}^2 > V,$$

$$f(G_n) = G_n - g_n^2 > V.$$

Noting that $f(\bar{G})$ is cancave in $\bar{G}$, we have that $\bar{G}-\bar{g}^2 > V$. This is a contradiction.

We define

$$S = \{ n \mid V_n \leqq V,\ n=0,1,\ldots,N \}.$$

If $S = \phi$, then lemmas 3 and 4 imply that our problem P_V has no feasible policy. Otherwise, we let

$$K+1 = \min\{n \mid n \in S\},$$

that is, $V_n > V$, $n=0,1,\ldots,K$, $V_{K+1} \leqq V$. Then, from lemma 4, it follows that

$$\bar{G} \leqq G_K. \tag{5.4}$$

Furthermore, the following equation (5.5) with respect to $\alpha \in [0,1)$ is easily seen to have a unique solution $\bar{\alpha}$.

$$\alpha G_K+(1-\alpha)G_{K+1} - \{\alpha g_K+(1-\alpha)g_{K+1}\}^2 = V. \tag{5.5}$$

Using $\bar{\alpha}$, we make the randomized policy $A_{K,K+1}$ from policies A_K and A_{K+1} as follows.

$$d_i^{A_K} = \bar{\alpha}\pi_i^{A_K}/\{\bar{\alpha}\pi_i^{A_K} + (1-\bar{\alpha})\pi_i^{A_{K+1}}\},\quad d_i^{A_{K+1}} = 1 - d_i^{A_K}, \tag{5.6}$$

that is, under policy $A_{K,K+1}$ we choose at state i the actions which construct policies A_K and A_{K+1} with probabilities $d_i^{A_K}$ and $d_i^{A_{K+1}}$, respectively. We let $G_{K,K+1}$ and $g_{K,K+1}$ denote the square mean and the mean for policy $A_{K,K+1}$, then $g_{K,K+1}$ is the solution of the following equation with respect to $g_{K,K+1}$, u_0, $u_1,\ldots,u_m$:

$$g_{K,K+1}+u_i = r_i^{A_K}d_i^{A_K}+r_i^{A_{K+1}}d_i^{A_{K+1}} + \sum_j (d_i^{A_K}p_{ij}^{A_K}+d_i^{A_{K+1}}p_{ij}^{A_{K+1}})u_j,\quad i \in I, \tag{5.7}$$

$$u_0 = 0.$$

From equations (5.6) and (5.7), we have that

$$\sum_i \{\bar{\alpha}\pi_i^{A_K}+(1-\bar{\alpha})\pi_i^{A_{K+1}}\}(g_{K,K+1}+u_i) \tag{5.8}$$

$$= \sum_i \{\bar{\alpha} r_i^{A_K} \pi_i^{A_K} + (1-\bar{\alpha}) r_i^{A_{K+1}} \pi_i^{A_{K+1}} + \bar{\alpha} \sum_j \pi_i^{A_K} p_{ij}^{A_K} u_j + (1-\bar{\alpha}) \sum_j \pi_i^{A_{K+1}} p_{ij}^{A_{K+1}} u_j\},$$

which implies that

$$(5.9) \qquad g_{K,K+1} = \bar{\alpha} g_K + (1-\bar{\alpha}) g_{K+1}.$$

In a similar way, we have that

$$(5.10) \qquad G_{K,K+1} = \bar{\alpha} G_K + (1-\bar{\alpha}) G_{K+1}.$$

Here, it should be noted that $G_{K,K+1} - g_{K,K+1}^2 = V$, that is, policy $A_{K,K+1}$ is a feasible policy and that $G_{K+1} \leqq G_{K,K+1} \leqq G_K$. Furthermore, from equations (5.3) and (5.5), it holds that

$$(5.11) \qquad \theta_{K+1} G_{K,K+1} - g_{K,K+1} = \theta_{K+1} G_K - g_K = \theta_{K+1} G_{K+1} - g_{K+1} \leqq \theta_{K+1} \bar{G} - \bar{g}.$$

As for $G_{K,K+1}$ and $\bar{G}$, we have the following lemma.

Lemma 5. It holds that $\bar{G} \leqq G_{K,K+1}$.

Proof: From equation (5.11), we have that

$$\bar{G} - \bar{g}^2 \geqq h(\bar{G}) \equiv \bar{G} - \{\theta_{K+1}(\bar{G} - G_{K+1}) + g_{K+1}\}^2,$$

$$h(G_K) = G_K - g_K^2 > V,$$

$$h(G_{K,K+1}) = G_{K,K+1} - g_{K,K+1}^2 = V.$$

Noting that $h(\bar{G})$ is cancave, it is easily seen that $\bar{G} \leqq G_{K,K+1}$.

From equation (5.11) and lemma 5, we have that

$$(5.12) \qquad g_{K,K+1} \geqq \theta_{K+1}(G_{K,K+1} - \bar{G}) + \bar{g} \geqq \bar{g},$$

which implies that policy $A_{K,K+1}$ is an optimal policy.

To conclude this section, we give the following theorem which is the result of the above discussion.

Theorem 1.

i) If $V_0 \leqq V$, then policy A_0 is optimal.

ii) If $S = \phi$, then problem P_V has no feasible policy.

iii) Otherwise, then policy $A_{K,K+1}$ is optimal.

It should be noted that there exists an optimal policy which is a mixture of at most two pure policies.

6. Sequential replacement problem

As an example of a variance constrained MDP, we consider a discrete time Markovian deterioration system whose level of function is quantized in four states 0,1,2,3 in the order of increasing deterioration (for detail, see Kawai [4]). The state 0 is a good state, i.e., the system is like new, the states 1 and 2 are detrioration states and the state 3 is a failed state. When the system is in deterioration states, we have two actions, i.e., to continue operation without replacement (action 1) and to replace by a new one (action 2). When the system is in the good state, we continue to operate and when it fails, we must replace. We assume that it takes one period to replace. Then, in our MDP, the state space and the action space are as follows:

$$I = \{0,1,2,3\},$$

$$K_0 = \{1\},\quad K_1 = \{1,2\},\quad K_2 = \{1,2\},\quad K_3 = \{2\}.$$

The one-step transition probabilities are assumed to be

$$p^1_{00}=p^1_{01}=p^1_{11}=p^1_{12}=p^1_{22}=p^1_{23}=0.5,\quad p^2_{10}=p^2_{20}=p^2_{30}=1.$$

As for the cost, we take into consideration only the replacement cost, i.e., the preventive replacement cost is 10 and the corrective replacement cost is 13. Then, our MDP has the following reward structure:

$$r^1_{00}=r^1_{01}=r^1_{11}=r^1_{12}=r^1_{22}=r^1_{23}=0-(-13)=13,\quad r^2_{10}=r^2_{20}=-10-(-13)=3,\quad r^2_{30}=-13-(-13)=0.$$

For this MDP, we have four pure policies $P^3=(1,1,1,2)$, $P^2=(1,1,2,2)$, $P^1=(1,2,1,2)$ and $P^4=(1,2,2,2)$. It should be noted that we need not distinguish P^1 from P^4, since they give the same system behavior.

Solving the parametric MDP corresponding to our variance constrained problem, we have that

$$N=2,\quad \theta_1=1/55,\ \theta_2=1/16,\quad A_0=P^3,\ A_1=P^2,\ A_2=P^1.$$

The mean and the variance for each policy A_n, $n=0,1,2$ is derived from equations (2.2)-(2.4) as follows.

$$g_0=78/7,\ g_1=11,\ g_2=29/3,\quad V_0=1014/49,\ V_1=16,\ V_2=200/9.$$

Comparing the variances, we have that

$$V_1 < V_0 < V_2.$$

Hence, we have three cases as for the upper bound of the variance V from theorem 1.

i) If $V \geqq 1014/49(=V_0)$, then policy $A_0(=P^3)$ is optimal.

ii) If $(V_1=)16 \leqq V < 1014/49$, then policy $A_{0,1}$, i.e., the mixture of two policies P^3 and P^2, is optimal. From equations (5.5) and (5.6) optimal probabilities to choose each action at deterioration states are given by

$d_1^1=1$ (both policies P^2 and P^3 select action 1 at state 1),

$d_2^1=10\xi/(7+3\xi)$, $d_2^2= 1-d_2^1$, where $\xi=7(33-\sqrt{1153-4V})/2$.

iii) Otherwise, there is no feasible policy.

References

[1] Benito, F.: Calculating the variance in Markov-processes with random reward. *Trabajos de Estadistica y de Investigacion Operativa*, Vol.33, (1892), 73-85.

[2] Derman, C.: *Finite state Markovian decision processes*. Academic Press, New York, 1970.

[3] Howard, R.: *Dynamic programming and Markov processes*. MIT Press, Cambridge, 1960.

[4] Kawai, H.: An optimal ordering and replacement policy of a Markovian deterioration system under incomplete observation part II. *Journal of the oprations Research Society of Japan*, Vol.26(1983), 293-308.

[5] Mandl, P.: On the variance of controlled Markov chain. *Kybernetika*, Vol.7 (1971), 1-12.

[6] Sobel, M. J.: The variance of discounted Markov decision processes. *Journal of Applied Probability*, Vol.19 (1982), 794-802.

[7] White, D. J.: Dynamic programming, Markov chains and the method of successive approximations. *Journal of Mathematical Analysis and Applications*, Vol.6 (1963), 654-664.

Hajime KAWAI: Department of Business Administration, School of Economics, University of Osaka Prefecture, Sakai, Osaka, 591, Japan.