Ann. Inst. Statist. Math.
35 (1983), Part A, 185-191

# ON AN APPROXIMATION FOR A MULTI-STAGE DECISION PROBLEM

SHINTARO SONO

(Received Apr. 27, 1982; revised Dec. 6, 1982)

## Summary

An approximation procedure based on conditional expectation and minimization operations is considered for a multi-stage decision problem from mathematical view point. This approximation procedure is applicable to adaptive decision processes, i.e., decision processes including unknown parameters.

## 1. Introduction

Consider the model,

$$x_{i+1}=A_i x_i+B_i u_i+q_i\,, \qquad i=0, \cdots, N-1\,, \tag{1.1}$$

where $N$ is a positive integer, all $A$'s and $B$'s are $(n, n)$- and $(n, r)$-known constant real valued matrices, respectively, $x_0$ and $q$'s are $n$-dimensional real valued random vectors on some probability space, $(W, F, P)$, and each $u_i$ is a function of $(x_j)_{j=0}^{i}$ which may depend on the selection of $(u_j)_{j=0}^{i-1}$. Each $x_i$, $u_i$, and $q_i$ are called the state, the decision, and the plant noise at the $i$th stage, respectively. It is assumed that each $x_i$ is accurately observed at the $i$th stage, $i=0, \cdots, N-1$, $x_0$ and all $q$'s have finite variance matrices and the probability distribution of $(x_0, (q_i)_{i=0}^{N-1})$, $P_{(x_0, (q_i)_{i=0}^{N-1})}$, is independent of the choice of $(u_i)_{i=0}^{N-1}$. The loss function is defined as

$$J:=\sum_{i=1}^{N} W_i\,, \tag{1.2}$$

where $W_i$ is the loss function of the decision at the $(i-1)$th stage, which depends only on $(x_i, u_{i-1})$, written as $W_i=W_i(x_i, u_{i-1})$. It is well-known that the multi-stage decision problem of finding $u$'s which minimize the expected loss, $\mathrm{E}(J|x_0)$, under the given distribution, $P_{(x_0, (q_i)_{i=0}^{N-1})}$, is solved by the backward induction or by the DP-algorithm, i.e., by the functional equation,

$$(1.3)\qquad \gamma_i^* = \min\{\mathrm{E}(W_i + \gamma_{i+1}^* | x^{i-1});\ u_{i-1} \in R^r\} = \mathrm{E}(W_i + \gamma_{i+1}^* | x^{i-1})|_{u_{i-1}=u_{i-1}^*},$$

where $x^{i-1} := (x_j)_{j=0}^{i-1}$ and $i=1, \cdots, N$ and $\gamma_{N+1} := 0$. The minimization in (1.3) is carried out for any given $x^{i-1}$ and $u^{i-2} := (u_j)_{j=0}^{i-2}$, $u^{-1} := 0$, and $u_{i-1}^*$ is one of the values which minimize the conditional expected loss,

$$(1.4)\qquad \gamma_i := \mathrm{E}(W_i + \gamma_{i+1}^* | x^{i-1}),$$

where $\gamma_i$ depends on $x^{i-1}$ and the choice of $u^{i-2}$. For applications of (1.3) see, for example, Aoki [1], Chapter I–Chapter III, or for concise exposition, see, for example, Suzuki [4], Chapter 6 (In general the sequence of the symbols, $(z_j)_{j=0}^i$, is written $z^i$, i.e., $z^i := (z_j)_{j=0}^i$, and put $z^{-1} := 0$.).

Consider any random variable, $\theta$, defined on $(W, F, P)$, such that $P_{(x_0, q^{N-1}, \theta)}$ is independent of the choice of $u^{N-1}$ and the functional equation,

$$(1.5)\qquad \gamma_{i,\theta}^* = \min\{\gamma_{i,\theta};\ u_{i-1} \in R^r\} = \gamma_{i,\theta}|_{u_{i-1}=u_{i-1,\theta}^*}, \qquad i=1, \cdots, N,$$

where $\gamma_{N+1,\theta} := 0$ and

$$(1.6)\qquad \gamma_{i,\theta} := \mathrm{E}(W_i + \gamma_{i+1,\theta}^* | \theta, x^{i-1})$$

($u_{i-1,\theta}^*$ is one of the values which minimize (1.6).)

Remark that if $\theta \equiv 0$, then (1.5) is identical with (1.3) in the sense of "almost surely". M. Aoki considers the approximation of (1.3) by the expectations of (1.5), using the posterior distributions of $\theta$, under some assumptions, the range of $\theta$ is a finite set, the plant noises, given $\theta$, are independent, etc. (see Aoki [1], pp. 224–241). In the following discussions the approximation method is extended to the case such that $\theta$ is arbitrary and the plant noises, given $\theta$, may be dependent. The approximation error is estimated by some procedure. In some special cases the procedure becomes rather simple.

Unless otherwise stated, all vectors and matrices are real valued and $\langle\!\langle a, b\rangle\!\rangle_M := a'Mb$, $\|a\|_M^2 := \langle\!\langle a, a\rangle\!\rangle_M$, $\mathrm{V}(x, y|z) := \mathrm{E}(xy'|z) - \mathrm{E}(x|z)\,\mathrm{E}(y|z)'$, $\mathrm{V}(x|z) := \mathrm{V}(x, x|z)$, where $a$, $b$ and $M$ are vectors and a matrix, respectively, and $x$, $y$ and $z$ are random vectors and a random variable, respectively. Any generalized inverse of a matrix, $M$, is written $M^-$. For elementary properties of generalized inverses see, for example, Iri and Kan [2], Chapter 8, or Rao [3], Chapter I.

## 2. Approximation and error estimation

Consider the quantities,

$$(2.1)\qquad \tilde{\gamma}_i^* := \min\{\mathrm{E}(\gamma_{i,\theta} | x^{i-1});\ u_{i-1} \in R^r\} \quad \text{and}$$

(2.2) $$\hat{\gamma}_i^* := \mathrm{E}\,(\gamma_{i,\theta}^* | x^{i-1})\,, \qquad i=1, \cdots, N+1\,.$$

These quantities are the approximations of $\gamma_i^*$ and the errors of these approximations are defined as

(2.3) $$\tilde{\Delta}\gamma_i := \gamma_i^* - \tilde{\gamma}_i^*\,,$$

(2.4) $$\Delta\gamma_i := \gamma_i^* - \hat{\gamma}_i^*\,, \qquad i=1, \cdots, N+1\,,$$

and the difference between (2.1) and (2.2) is defined by

(2.5) $$\delta\gamma_i := \tilde{\gamma}_i^* - \hat{\gamma}_i^*\,, \qquad i=1, \cdots, N+1\,.$$

It is clear from the definitions that (2.3), (2.4) and (2.5) are non-negative and

(2.6) $$\tilde{\Delta}\gamma_i = \gamma_i^* - \tilde{\gamma}_i^* = \Delta\gamma_i - \delta\gamma_i\,, \qquad i=1, \cdots, N+1\,.$$

For estimation of $\Delta\gamma_i$, $i=1, \cdots, N+1$, the assumption (2.7) is used:

(2.7) $\mathrm{E}\,(\delta\gamma_i | x^{i-2-k})$ is independent of the choice of the decisions, $(u_j)_{j=i-2-k}^{N-1}$, for all $k=0, \cdots, i-1$, and for each $i=1, \cdots, N+1$.

Under the assumption (2.7) the following proposition (2.8) and the formula, (2.9), are obtained:

(2.8) $\mathrm{E}\,(\Delta\gamma_i | x^{i-2-k})$ is independent of the choice of the decisions, $(u_j)_{j=i-2-k}^{N-1}$, for all $k=0, \cdots, i-1$, and for each $i=1, \cdots, N+1$,

and

(2.9) $$\begin{aligned}\Delta\gamma_i &= \delta\gamma_i + \mathrm{E}\,(\Delta\gamma_{i+1} | x^{i-1}) \\ &= \delta\gamma_i + \mathrm{E}\left(\sum_{j=i+1}^{N} \delta\gamma_j | x^{i-1}\right), \qquad \text{for each } i=1, \cdots, N+1\,.\end{aligned}$$

(2.8) and (2.9) are established by the backward induction as the following:

For $i=N+1$ (2.8) and (2.9) are clear because $\gamma_{N+1}=0$ and $\gamma_{N+1,\theta}=0$.

For general case assume that (2.8) and (2.9) are true for all $i=j+1, \cdots, N+1$, for any fixed $j \leqq N$.

Then

$$\begin{aligned}\gamma_j^* &= \min_{u_{j-1}} \mathrm{E}\,(W_j + \gamma_{j+1}^* | x^{j-1}) = \min_{u_{j-1}} \mathrm{E}\,(W_j + \hat{\gamma}_{j+1}^* + \Delta\gamma_{j+1} | x^{j-1}) \\ &= \min_{u_{j-1}} \{\mathrm{E}\,(W_j + \gamma_{j+1,\theta}^* | x^{j-1}) + \mathrm{E}\,(\Delta\gamma_{j+1} | x^{j-1})\}\end{aligned}$$

$$=\tilde{\gamma}_j^*+\mathrm{E}\,(\Delta\gamma_{j+1}|x^{j-1})\,,$$

therefore, using (2.6),

$$\tilde{\Delta}\gamma_j=\Delta\gamma_j-\delta\gamma_j=\mathrm{E}\,(\Delta\gamma_{j+1}|x^{j-1})\,,\quad \text{i.e.,}$$

$$\Delta\gamma_j=\delta\gamma_j+\mathrm{E}\,(\Delta\gamma_{j+1}|x^{j-1})\,.$$

Therefore, from the assumption (2.7) and the induction's assumption (2.8) and (2.9) are established for $i=j$.
(In the above discussion the well-known formula,

$$\mathrm{E}\,(\mathrm{E}\,(x|\theta,\,y)|y)=\mathrm{E}\,(x|y)\,,$$

where $y$ and $\theta$ are random variables on $(W, F, P)$ and $x$ is a random vector on $(W, F, P)$, is used freely.)

## 3. Quadratic loss function

Consider the special case,

$$W_i=W_i(x_i,\,u_{i-1})=\|x_i\|^2_{V_i}+\|u_{i-1}\|^2_{P_{i-1}}\,,\qquad i=1,\,\cdots,\,N\,, \tag{3.1}$$

where $V_i$ and $P_{i-1}$ are known constant real valued symmetric nonnegative definite $(n, n)$ and $(r, r)$ matrices, respectively. Then the functional equation (1.5) is solved by the backward induction as the following:

For $i=N$ from (1.5) and (1.6)

$$\gamma_{N,\theta}=\|u_{N-1}-u^*_{N-1,\theta}\|^2_{S_{N-1}}+\gamma^*_{N,\theta}\,, \tag{3.2}$$

where

$$u^*_{N-1,\theta}:=-S^-_{N-1}B'_{N-1}V_N(A_{N-1}x_{N-1}+m_{N-1,\theta})\,, \tag{3.3}$$

$$S_{N-1}:=P_{N-1}+B'_{N-1}V_NB_{N-1}\,, \tag{3.4}$$

$$m_{N-1,\theta}:=\mathrm{E}\,(q_{N-1}|\theta,\,x^{N-1})\,, \tag{3.5}$$

$$\gamma^*_{N,\theta}=\|A_{N-1}x_{N-1}+m_{N-1,\theta}\|^2_{J_{N-1}}+\mathrm{tr}\,(V_N\Sigma_{N-1,\theta})\,, \tag{3.6}$$

$$J_{N-1}:=V_N-V_NB_{N-1}(P_{N-1}+B'_{N-1}V_NB_{N-1})^-B'_{N-1}V_N\,, \tag{3.7}$$

$$\Sigma_{N-1,\theta}:=V(q_{N-1}|\theta,\,x^{N-1})\,. \tag{3.8}$$

(Remark that $J_{N-1}$ is symmetric nonnegative definite.)

For general case assume that

$$\gamma_{i,\theta}=\|u_{i-1}-u^*_{i-1,\theta}\|^2_{S_{i-1}}+\gamma^*_{i,\theta}\,, \tag{3.9}$$

where

$$u^*_{i-1,\theta} := -S^-_{i-1}B'_{i-1}V^{(i)}_i(A_{i-1}x_{i-1}+m^{(i-1)}_{i-1,\theta}) ,$$

$$\gamma^*_{i,\theta} = \|A_{i-1}x_{i-1}+m^{(i-1)}_{i-1,\theta}\|^2_{J_{i-1}}+R_{i,\theta} ,$$

$$S_{i-1} := P_{i-1}+B'_{i-1}V^{(i)}_iB_{i-1} ,$$

$$J_{i-1} := V^{(i)}_i - V^{(i)}_iB_{i-1}S^-_{i-1}B'_{i-1}V^{(i)}_i ,$$

and $V^{(i)}_i$ is symmetric nonnegative definite (therefore $J_{i-1}$ is symmetric nonnegative definite) and $\mathrm{E}\,(m^{(i-1)}_{i-1,\theta}|\theta, x^{i-1-k})$ and $\mathrm{E}\,(R_{i,\theta}|\theta, x^{i-1-k})$ are independent of the choice of the decisions, $(u_j\,;\ j\geqq i-1-k)$, for all $k=0, \cdots, i$.

Then

$$\gamma_{i-1,\theta} = \|u_{i-2}-u^*_{i-2,\theta}\|^2_{S_{i-2}}+\gamma^*_{i-1,\theta} , \tag{3.10}$$

where

$$u^*_{i-2,\theta} := -S^-_{i-2}B'_{i-2}V^{(i-1)}_{i-1}(A_{i-2}x_{i-2}+m^{(i-2)}_{i-2,\theta}) ,$$

$$S_{i-2} := P_{i-2}+B'_{i-2}V^{(i-1)}_{i-1}B_{i-2} ,$$

$$V^{(i-1)}_{i-1} := V_{i-1}+A'_{i-1}J_{i-1}A_{i-1} ,$$

$$m^{(i-2)}_{i-2,\theta} := m_{i-2,\theta}+(V^{(i-1)}_{i-1})^-A'_{i-1}J_{i-1}m^{(i-1)}_{i-1,i-2,\theta} ,$$

$$m^{(i-1)}_{i-1,i-2,\theta} := \mathrm{E}\,(m^{(i-1)}_{i-1,\theta}|\theta, x^{i-2}) ,$$

$$\gamma^*_{i-1,\theta} = \|A_{i-2}x_{i-2}+m^{(i-2)}_{i-2,\theta}\|^2_{J_{i-2}}+R_{i-1,\theta} ,$$

$$J_{i-2} := V^{(i-1)}_{i-1} - V^{(i-1)}_{i-1}B_{i-2}S^-_{i-2}B'_{i-2}V^{(i-1)}_{i-1} ,$$

$$\begin{aligned} R_{i-1,\theta} := {} & \|m^{(i-1)}_{i-1,i-2,\theta}\|^2_{J^{(i-1)}_{i-1}}+\mathrm{tr}\,(V^{(i-1)}_{i-1}\Sigma_{i-2,\theta}) \\ & +\mathrm{E}\,(R_{i,\theta}|\theta, x^{i-2})+\mathrm{tr}\,(J_{i-1}\Sigma^{(i-1)}_{(i-2),\theta}) , \end{aligned}$$

$$J^{(i-1)}_{i-1} := J_{i-1}-J_{i-1}A_{i-1}(V^{(i-1)}_{i-1})^-A'_{i-1}J_{i-1} ,$$

$$\Sigma_{i-2,\theta} := V(q_{i-2}|\theta, x^{i-2}) ,$$

$$\Sigma^{(i-1)}_{(i-2),\theta} := V(A_{i-1}q_{i-2}+m^{(i-1)}_{i-1,\theta},\ m^{(i-1)}_{i-1,\theta}|\theta, x^{i-2}) ,$$

and $J_{i-2}$ and $J^{(i-1)}_{i-1}$ are symmetric nonnegative definite and $\mathrm{E}\,(m^{(i-2)}_{i-2,\theta}|\theta, x^{i-2-k})$ and $\mathrm{E}\,(R_{i-1,\theta}|\theta, x^{i-2-k})$ are independent of the choice of the decisions, $(u_j\,;\ j\geqq i-2-k)$, for all $k=0, \cdots, i-1$.

(Put $\gamma_{N+1,\theta}:=0,\ m^{(N)}_{N,\theta}:=0,\ R_{N+1,\theta}:=0,\ V^{(N+1)}_{N+1}:=0,\ J_N:=0.$)

Especially if $\theta\equiv 0$, then the functional equation, (1.3), is solved recursively by (3.9) and (3.10).

From (3.9) and (3.10) we have

$$\begin{aligned} u^*_{i-1} &= \mathrm{E}\,(u^*_{i-1,\theta}|x^{i-1}) \\ &= -S^-_{i-1}B'_{i-1}V^{(i)}_i(A_{i-1}x_{i-1}+\mathrm{E}\,(m^{(i-1)}_{i-1,\theta}|x^{i-1})) , \qquad i=1, \cdots, N , \end{aligned} \tag{3.11}$$

and, from (2.5), we have

$$\delta\gamma_i = \mathrm{E}\,(\|u^*_{i-1,\theta} - \mathrm{E}\,(u^*_{i-1,\theta}|x^{i-1})\|^2_{S_{i-1}}|x^{i-1}) \tag{3.12}$$
$$= \mathrm{E}\,(\|m^{(i-1)}_{i-1,\theta} - \mathrm{E}\,(m^{(i-1)}_{i-1,\theta}|x^{i-1})\|^2_{C_{i-1}}|x^{i-1})$$
$$= \mathrm{tr}\,(C_{i-1}\,\mathrm{V}\,(m^{(i-1)}_{i-1,\theta}|x^{i-1}))\,, \qquad i=1, \cdots, N\,,$$

where

$$C_{i-1} := V_i^{(i)} B_{i-1} S_{i-1}^{-} B'_{i-1} V_i^{(i)}\,. \tag{3.13}$$

Therefore, using (2.9) and (3.12), the error estimation of the approximation, (2.2), is accomplished because $P_{(x_0, q^{N-1}, \theta)}$ is independent of the choice of the decisions, $u^{N-1}$.

Consider the special case; $q$'s, given $\theta$, are independently identically distributed random variables, $q_i|\theta \sim N_n(\theta, \Sigma)$, $i=0, \cdots, N-1$, $\theta \sim N_n(\theta_0, \Sigma_0)$, and $(q^{N-1}, \theta)$ and $x_0$ are stochastically independent. $\Sigma$ and $\Sigma_0$ are positive definite, each $B_{i-1}$ is equal to the unit matrix of order $n$, $i=1, \cdots, N$, and $P_{i-1}=0$, $i=1, \cdots, N$. Then

$$\theta|(x_0, q^{i-1}) \sim N_n(\theta_{(i)}, \Sigma_{(i)})\,, \qquad i=0, \cdots, N-1\,,$$

where $\Sigma_{(i)} := (\Sigma_0^{-1} + i\Sigma^{-1})^{-1}$, $i=0, \cdots, N-1$, and

$$\theta_{(i)} := \Sigma_{(i)}(\Sigma_0^{-1}\theta_0 + i\Sigma^{-1}\bar{q}_{(i-1)})\,, \qquad i=0, \cdots, N-1\,,$$

where

$$\bar{q}_{(i-1)} := \left(\sum_{j=0}^{i-1} q_j\right)\Big/ i\,, \qquad i=1, \cdots, N-1\,, \text{ and } \bar{q}_{(0)} := 0\,.$$

(Remark that $(x_0, q^{i-1})$ is uniquely determined by $x^i$.)

Therefore

$$\delta\gamma_i = \mathrm{tr}\,(V_i(\Sigma_0^{-1} + (i-1)\Sigma^{-1})^{-1})\,, \qquad i=1, \cdots, N\,,$$

$$\gamma^*_{i,\theta} = \mathrm{tr}\left(\left(\sum_{j=i}^{N} V_j\right)\Sigma\right), \qquad i=1, \cdots, N\,, \quad \text{and}$$

$$\hat{\gamma}_i^* = \mathrm{E}\,(\gamma^*_{i,\theta}|x^{i-1}) = \mathrm{tr}\left(\left(\sum_{j=i}^{N} V_j\right)\Sigma\right), \qquad i=1, \cdots, N\,.$$

From these results

$$(i-1)^{-1}\,\mathrm{tr}\,(V_i(\Sigma_0^{-1} + \Sigma^{-1})^{-1}) \leqq \delta\gamma_i \leqq (i-1)^{-1}\,\mathrm{tr}\,(V_i\Sigma)\,, \qquad i=2, \cdots, N\,,$$

and $\delta\gamma_1 = \mathrm{tr}\,(V_1\Sigma_0)$, and

$$\underline{c}(\Sigma) \sum_{j=i}^{N} \mathrm{tr}\,(V_j) \leqq \hat{\gamma}_i^* \leqq \bar{c}(\Sigma) \sum_{j=i}^{N} \mathrm{tr}\,(V_j)\,, \qquad i=1, \cdots, N\,,$$

where $\underline{c}(\Sigma)$ and $\bar{c}(\Sigma)$ are minimum and maximum eigenvalues of $\Sigma$, respectively. Therefore by application of these inequalities $\Delta\gamma_i$, $i=1, \cdots, N$, etc., are estimated, for example,

$$\frac{\tilde{\Delta}\gamma_1}{\hat{\gamma}_1^*}=\frac{\Delta\gamma_1-\delta\gamma_1}{\hat{\gamma}_1^*}=\frac{\sum_{j=2}^{N}\delta\gamma_j}{\hat{\gamma}_1^*}$$
$$\leqq\frac{\bar{c}(\Sigma)\sum_{i=2}^{N}(i-1)^{-1}\operatorname{tr}(V_i)}{\underline{c}(\Sigma)\sum_{j=1}^{N}\operatorname{tr}(V_j)}$$
$$\leqq K_N\cdot\left(\sum_{k=1}^{N-1}k^{-1}\right)\Big/N=K_N\cdot\left(\log(N-1)+C-\frac{1}{2(N-1)}+O\left(\frac{1}{(N-1)^2}\right)\right)\Big/N\,,$$
$$N\to\infty\,,$$

where $K_N:=\dfrac{\bar{c}(\Sigma)\max(\operatorname{tr}(V_i)\,;\ i=2, \cdots, N)}{\underline{c}(\Sigma)\min(\operatorname{tr}(V_i)\,;\ i=1, \cdots, N)}$ and $C$ is Euler's constant, and each $V_i$, $i=1, \cdots, N$, is assumed to be a non-zero matrix. In the above special case if $V_i=0$, $i=1, \cdots, N-1$, and $V_N$ is non-zero, then $\delta\gamma_i=0$, $i=1, \cdots, N-1$, $\delta\gamma_N=\operatorname{tr}(V_N(\Sigma_0^{-1}+(N-1)\Sigma^{-1})^{-1})$, $\hat{\gamma}_1^*=\operatorname{tr}(V_N\Sigma)$, $\tilde{\Delta}\gamma_1=\Delta\gamma_1=\delta\gamma_N$.

## Acknowledgement

The writer wishes to thank Professor Y. Suzuki and the referee for critical and valuable comments.

GRADUATE SCHOOL OF TOKYO UNIVERSITY

## REFERENCES

[1] Aoki, M. (1967). *Optimization of Stochastic Systems*, Academic Press, New York, 224-241.
[2] Iri, M. and Kan, T. (1977). *Linear Algebra—Matrices and Normal Forms* (in Japanese), Kyôiku Shuppan, Tokyo, Chapter 8.
[3] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd edition, John Wiley & Sons, New York, Chapter I.
[4] Suzuki, Y. (1978). *Statistical Analysis* (in Japanese), Chikuma Shobo, Tokyo, Chapter 6.