

# ベイズの逐次バッチサイズ決定問題について

兵庫県立大学 経営学部 濱田年男 (Toshio Hamada)

School of Business Administration,

University of Hyogo

## 1 緒言

ベイズの逐次バッチサイズ決定問題とは以下のような問題であるとする。1台の機械を用いて同一種類の  $n$  個のジョブを処理する。ジョブの処理時間は一定であり、その値は既知である。 $n$  個のジョブは複数個のバッチに分割され、個々のバッチごとに処理されるものとし、その大きさをバッチサイズと呼ぶことにする。バッチ内の最初のジョブを処理する前に準備時間 (setup time) が必要であり、バッチ内の2番目以降のジョブは、1番目のジョブに引き続いて加工されるものとする。同一バッチ内のすべてのジョブが処理されるまで、バッチ内のすべてのジョブはその機械に留まらなければならないものとする。したがって同一バッチ内のすべてのジョブの完了時刻は同じである。処理時間はすべて同じで既知であるが、準備時間は確率変数であり、その1つのパラメータの値は未知で、事前分布が与えられているものとする。目的はすべてのジョブの完了時間の総和を最小にすることである。そのためにはバッチサイズをどのように決めればよいかを決定する問題である。

一般にスケジューリング問題においては、加工を行う前の準備時間は処理時間に含めて取り扱われてきていた。これは個々のジョブを単独に取り扱うために、そのジョブの処理を行うのに必要な準備時間と処理時間は、共に同じ1台の機械を占有するために、まとめて処理時間として取り扱うことができたからである。しかし複数のジョブを1つのバッチとして処理する場合に、そのバッチ内のすべてのジョブに共通な準備作業を行うための準備時間を一度だけ考慮し、複数のジョブの処理時間だけを分離して取り扱う場合には、バッチサイズが大きいか小さいかによって、ジョブの完了時間 (時刻0で加工開始可能であるとすれば滞留時間と同等である) の総和に変化が生じてくる。このような理由のために、バッチサイズの決定は1つの重要な意思決定の要因になる。

スケジューリング問題において、ジョブのバッチサイズを考慮した研究は、1985年頃から多くの研究がなされてきており、その中でも特に単一種類のジョブを引き続いて複数個加工を行う問題は、Santos and Magazine [5], Naddef and Santos [4], Shallcross [6] 等において研究されてきた。これらにおいては準備時間と処理時間は共に定数である。したがって、準備時間と処理時間が与えられれば、バッチサイズが決定できる。

処理時間は機械の性能等により決定されるものであるが、準備時間は人間が関与する可能性が高く、定数と考えるよりは確率変数と考えることができる場合がある。もしその分布に関する情報が不十分であれば、学習を行うことにより、情報を収集しながら決定を行うという逐次決定問題となる。

このような問題の中で、処理時間がパラメータ  $u$  の指数分布に従う場合には、ジョブ数が4以下の問題が柳井 [7] において解かれている。また一般の  $n$  についての最適解の構造は濱田 [1] において与えられている。

本研究においては、バッチサイズを高々2までとすることにより、ジョブを1つずつ処理するか、あるいは2つをバッチとして処理するかという最も簡単な場合を考察する。このような問題を考察することは、他の逐次決定問題との比較において重要であると考えられる。

## 2 逐次バッチサイズ決定問題

1台の機械で  $n$  個のジョブを1個ずつ処理するものとする。ジョブの処理時間は一定であるととし、一般性を失うことなく1であるとする。また準備時間は処理時間に対する相対的な量として考え、その値はパラメータが  $u$  の指数分布に従う確率変数  $X$  であり、その確率密度関数は

$$\varphi(x|u) = \begin{cases} ue^{-ux}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0, \end{cases}$$

で与えられるものとする。ここでパラメータ  $u$  の値は未知であり、事前分布として、パラメータが  $(w, \alpha)$  のガンマ分布を仮定できるものとする。したがってその確率密度関数は

$$\psi(u|w, \alpha) = \begin{cases} \Gamma(\alpha)^{-1} w^\alpha u^{\alpha-1} e^{-wu}, & \text{if } u \geq 0, \\ 0, & \text{if } u < 0, \end{cases}$$

で与えられる。

ジョブは1つずつ処理する(バッチサイズ1)か、あるいは2つをバッチとして処理する(バッチサイズ2)かのいずれかであるとする。いずれの場合もその処理を行う前に準備時間が必要である。したがって前者の場合には準備時間  $X$  と処理時間1が生じ、後者の場合には、準備時間  $X$  と2つのジョブの処理時間の合計となる処理時間2が生じる。

残りのジョブ数が  $n$  で事前分布のパラメータが  $(w, \alpha)$  のときに、まずバッチサイズが1と2のいずれであるかの決定を行い、その後準備時間の観察値  $X = x$  を得た後に、パラメータ  $u$  の事前分布のパラメータを事後分布のパラメータに更新する。バッチサイズが1と2のいずれの決定を行っても、事後分布のパラメータはベイズの定理により  $(w+x, \alpha+1)$  となる。しかし残りのジョブ数は、バッチサイズ1という決定を行った場合には  $n-1$  となり、またバッチサイズ2という決定を行った場合には  $n-2$  となる。新たな残りのジョブ数に対してこれを繰り返していく。目的は  $n$  個のジョブの完了時刻の総和の期待値を最小にすることである。

## 3 動的計画法による定式化

各決定時点における状態は、残りのジョブ数  $n$  とそのときの事前分布のパラメータ  $w$  と  $\alpha$  からなるベクトル  $(n; w, \alpha)$  で与えられる。状態が  $(n; w, \alpha)$  においてバッチサイズ1として、準備時間  $X = x$  を得たとき、状態は  $(n-1; w+x, \alpha+1)$  となる。また状態が  $(n; w, \alpha)$  においてバッチサイズ2として、準備時間  $X = x$  を得たときには、状態は  $(n-2; w+x, \alpha+1)$  となる。

現在の状態が  $(n; w, \alpha)$  のときに、以後は最適に行動したときの完了時刻の総和の期待値を  $F_n(w, \alpha)$  とする。また現在の状態が  $(n; w, \alpha)$  のときに、まず1個あるいは2個のジョブを処理し、以後は最適に行動したときの完了時刻の総和の期待値をそれぞれ  $F_n^1(w, \alpha)$ , および  $F_n^2(w, \alpha)$  とする。このとき

$$\begin{aligned} F_n(w, \alpha) &= \min \{F_n^1(w, \alpha), F_n^2(w, \alpha)\} \quad (n = 2, 3, 4, \dots) \\ F_1(w, \alpha) &= w(\alpha - 1)^{-1} \\ F_0(w, \alpha) &= 0 \end{aligned}$$

となる。ここに

$$\begin{aligned} F_n^1(w, \alpha) &= E_U[E_X[n(X+1) + F_{n-1}(w+X, \alpha+1)|U]|w, \alpha] \\ F_n^2(w, \alpha) &= E_U[E_X[n(X+2) + F_{n-2}(w+X, \alpha+1)|U]|w, \alpha] \end{aligned}$$

であり、

$$\begin{aligned} E_X[f(X)|U] &= \int_0^\infty f(x)ue^{-ux}dx \\ E_U[g(U)|w, \alpha] &= \int_0^\infty g(u)\frac{1}{\Gamma(\alpha)}w^\alpha u^{\alpha-1}e^{-wu}du \end{aligned}$$

である。

$$\begin{aligned} &E_U[E_X[f(X)|U]|w, \alpha] \\ &= \int_0^\infty \int_0^\infty f(x)ue^{-ux}dx \frac{1}{\Gamma(\alpha)}w^\alpha u^{\alpha-1}e^{-wu}du \\ &= \int_0^\infty f(x) \left\{ \int_0^\infty \frac{1}{\Gamma(\alpha+1)}(w+x)^{\alpha+1}u^\alpha e^{-(w+x)u}du \right\} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \frac{w^\alpha}{(w+x)^{\alpha+1}} dx \\ &= \int_0^\infty f(x) \frac{\alpha w^\alpha}{(w+x)^{\alpha+1}} dx \end{aligned}$$

以下においては、期待値  $E[f(X)|w, \alpha]$  は

$$E[f(X)|w, \alpha] = \int_0^\infty f(x) \frac{\alpha w^\alpha}{(w+x)^{\alpha+1}} dx$$

とする。このとき

$$\begin{aligned} F_n^1(w, \alpha) &= nw(\alpha - 1)^{-1} + n + E[F_{n-1}(w+X, \alpha+1)|w, \alpha] \\ F_n^2(w, \alpha) &= nw(\alpha - 1)^{-1} + 2n + E[F_{n-2}(w+X, \alpha+1)|w, \alpha] \end{aligned}$$

となる。

以下の等式が成立する。まず

$$F_n^1(w, \alpha) - F_n^2(w, \alpha) = -n + E[F_{n-1}(w+X, \alpha+1) - F_{n-2}(w+X, \alpha+1)|w, \alpha]$$

が得られる。また

$$\begin{aligned} F_n^1(w, \alpha) - F_{n-1}^1(w, \alpha) &= w(\alpha - 1)^{-1} + 1 \\ &\quad + E[F_{n-1}(w + X, \alpha + 1) - F_{n-2}(w + X, \alpha + 1) | w, \alpha] \\ F_n^2(w, \alpha) - F_{n-1}^2(w, \alpha) &= w(\alpha - 1)^{-1} + 2 \\ &\quad + E[F_{n-2}(w + X, \alpha + 1) - F_{n-3}(w + X, \alpha + 1) | w, \alpha] \\ F_n^2(w, \alpha) - F_{n-1}^1(w, \alpha) &= w(\alpha - 1)^{-1} + n + 1 \end{aligned}$$

が得られる。

特に  $n = 2$  に対しては,

$$F_2(w, \alpha) = \begin{cases} 3w(\alpha - 1)^{-1} + 3, & \text{if } 0 < w < r_2(\alpha) \\ 2w(\alpha - 1)^{-1} + 4, & \text{if } r_2(\alpha) \leq w \end{cases}$$

ここに  $r_2(\alpha) = \alpha - 1$  である。また

$$F_2^1(w, \alpha) = 3w(\alpha - 1)^{-1} + 3$$

$$F_2^2(w, \alpha) = 2w(\alpha - 1)^{-1} + 4$$

である。これらを用いて次の性質が得られる。

**補題 1**  $w > 0, \alpha > 1$  とする。  $f(x)$  が連続で狭義単調増加関数ならば,

$$E[f(X) | w, \alpha] = \int_0^\infty f(x) \frac{\alpha w^\alpha}{(w+x)^{\alpha+1}} dx$$

は  $w$  について連続で狭義単調増加,  $\alpha$  について連続で狭義単調減少である。

**補題 2**  $w > 0, \alpha > 1$  とする。このとき

(i)  $F_n^1(w, \alpha)$  および  $F_n^2(w, \alpha)$  は  $w$  について連続で狭義単調増加,  $\alpha$  について連続で狭義単調減少である。

(ii)  $F_n(w, \alpha)$  は  $w$  について連続で狭義単調増加,  $\alpha$  について連続で狭義単調減少である。

**補題 3**  $w > 0, \alpha > 1$  に対して

$$w(\alpha - 1)^{-1} + n \leq F_n(w, \alpha) - F_{n-1}(w, \alpha) \leq nw(\alpha - 1)^{-1} + n$$

が成立する。

**補題 4**  $n \geq 2, w > 0, \alpha > 1$  に対して

$$w(\alpha - 1)^{-1} - 1 \leq F_n^1(w, \alpha) - F_n^2(w, \alpha) \leq (n - 1)w(\alpha - 1)^{-1} - 1$$

が成立する。

これらの補題等を用いて次の定理が得られる.

**定理 1**  $n \geq 3$  に対して  $F_n^1(w, \alpha) - F_n^2(w, \alpha)$  は  $w$  について連続で狭義単調増加,  $\alpha$  について連続で狭義単調減少であり,  $w$  についての方程式  $F_n^1(w, \alpha) - F_n^2(w, \alpha) = 0$  は唯一の解  $r_n(\alpha)$  をもち,

$$F_n(w, \alpha) = \begin{cases} F_n^1(w, \alpha), & \text{if } 0 < w < r_n(\alpha) \\ F_n^2(w, \alpha), & \text{if } r_n(\alpha) \leq w \end{cases}$$

となる. さらに  $r_n(\alpha) \leq r_{n-1}(\alpha)$  であり, また  $r_n(\alpha) < r_{n-1}(\alpha + 1) < \dots < r_1(\alpha + n - 1)$  が成立する.

## 4 最適政策

定理 1 より,  $r_n(\alpha) < r_{n-1}(\alpha + 1) \leq r_{n-2}(\alpha + 1)$  が成立することは明らかである. したがって最適政策は以下ようになる.

### 最初のバッチサイズの決定:

今  $n$  個のジョブがあり, 現在の事前分布のパラメータが  $(w, \alpha)$  であるときに,

- (i) もし  $0 < w < r_n(\alpha)$  ならばバッチサイズ 1 として処理する. このとき, 残りのジョブ数は  $n - 1$  となる. また,
- (ii) もし  $r_n(\alpha) \leq w$  ならばバッチサイズ 2 として処理する. このとき, 残りのジョブ数は  $n - 2$  となる.

いずれの場合も準備時間の実現値  $x_1$  を観察して, 事後分布のパラメータは  $(w + x_1, \alpha + 1)$  となる.

### 2 番目のバッチサイズの決定:

- (I) 残りのジョブ数が  $n - 1$  のとき,
  - (i) もし  $0 < w + x_1 < r_{n-1}(\alpha + 1)$  ならばバッチサイズ 1 として処理する. このとき, 残りのジョブ数は  $n - 2$  となる. また,
  - (ii) もし  $r_{n-1}(\alpha + 1) \leq w + x_1$  ならばバッチサイズ 2 として処理する. このとき, 残りのジョブ数は  $n - 3$  となる.

いずれの場合も準備時間の実現値  $x_2$  を観察して, 事後分布のパラメータは  $(w + x_1 + x_2, \alpha + 2)$  となる.

- (II) 残りのジョブ数が  $n - 2$  のとき,
  - (i) もし  $0 < w + x_1 < r_{n-2}(\alpha + 1)$  ならばバッチサイズ 1 として処理する. このとき, 残りのジョブ数は  $n - 3$  となる. また,
  - (ii) もし  $r_{n-2}(\alpha + 1) \leq w + x_1$  ならばバッチサイズ 2 として処理する. このとき, 残りのジョブ数は  $n - 4$  となる.

いずれの場合も準備時間の実現値  $x_2$  を観察して, 事後分布のパラメータは  $(w + x_1 + x_2, \alpha + 2)$

となる。

3番目以降のバッチサイズの決定も同様である。

## 5 他の逐次決定問題との比較

他の学習を含む逐次決定問題としては、バンディット問題、単一機械スケジューリング問題があるが、本研究で扱っているモデルには大きな特徴がある。

バンディット問題の最も簡単なモデルとして、以下のようなものを考える。2つの実験  $a_0$  と  $a_1$  があり、毎回これらの2つの実験のいずれかを選択して、ひき続いて全部で  $n$  回の実験を行うものとする。実験  $a_0$  を行ったときにはパラメータが既知の分布から観察値を得る。また実験  $a_1$  を行ったときにはパラメータが未知の分布から観察値を得る。この未知パラメータについては、事前分布が与えられているものとする。目的は引き続いて行われる  $n$  回の実験から得られる観察値の和の期待値を最大にすることである。実験  $a_0$  および実験  $a_1$  を行った結果、観察値がそれぞれ区間  $(0, 1)$  および区間  $(0, u)$  上の一様分布から得られ、 $u$  の事前分布がパレート分布で与えられる場合は、Hamada[2]で論じられている。

単一機械スケジューリング問題の最も簡単なモデルとして、以下のようなものを考える。機械は1度に1個のジョブしか処理できないものとする。2種類のジョブ  $J_0$  と  $J_1$  がそれぞれ  $n_0$  個、 $n_1$  個存在する。ジョブ  $J_0$  の処理時間はパラメータが既知の分布に従い、 $J_1$  の処理時間はパラメータが未知の分布に従う。この未知パラメータについては、事前分布が与えられているものとする。これら  $n_0 + n_1$  個のジョブの完了時刻の総和の期待値を最小にするためには、ジョブをどのように処理していけばよいかを決定することである。この問題で処理時間が指数分布の場合には、Hamada[3]において論じられている。

バンディット問題では、実験  $a_0$  を行ったときには未知パラメータの事前分布は更新されず、実験  $a_1$  を行ったときのみ未知パラメータの事前分布が更新される。さらに実験  $a_0$  を行うのが最適ならば、以後は実験  $a_0$  だけを行うのが最適となり、問題は実験  $a_1$  の利用をいつ止めるかという最適停止問題に帰着する。したがって実験  $a_1$  を行う回数は未知であり、未知パラメータを持つ分布から得られる観察値の個数は未定であり、その点では本研究のバッチサイズ決定問題に類似する。

これに対してスケジューリング問題では、未知パラメータを持つ分布から得られる観察値の個数は一定である。この問題では、最初  $J_1$  が最適な場合に、 $J_1$  を行って観察値を得て、再び  $J_0$  と  $J_1$  のいずれが最適であるかを決定する。 $J_1$  が最適である限りこれを繰り返す。ある段階においてジョブ  $J_0$  が最適になれば以後ジョブ  $J_0$  がなくなるまでジョブ  $J_0$  を続けて処理し、ジョブ  $J_0$  がなくなった後はジョブ  $J_1$  だけが残るので、その後のジョブ  $J_1$  の処理によって得られる情報は最適な行動に対して何の影響も与えない。したがって最初  $J_1$  が最適な場合に、いつ  $J_0$  に切り替えるかが重要になり、この問題も最適停止問題に帰着する。

このようなバンディット問題とスケジューリング問題は共に最適停止問題に帰着する。これに対して、本研究のモデルでは、ある状態でバッチサイズ1が最適になっても、次の段階でバッチサイズ2が最適になることもあれば、逆のケースもあり、最適停止問題には帰着しない。これはバッチサイズが1と2のいずれであっても、必ず未知パラメータを持つ

分布から1つの観察値を得るからである。

謝辞 本研究は日本学術振興会の科学研究費補助金 (C)(2) 15510136 の補助により行われたものである。

#### 参考文献

- [1] 濱田 (2004) 学習型逐次バッチサイズ決定問題の最適政策について. 2004年度日本オペレーションズ・リサーチ学会秋季研究発表会アブストラクト集, 250-251.
- [2] Hamada, T. (1978) A uniform two-armed bandit problem: the parameter of one distribution is known. *Journal of the Japan Statistical Society* **8**, 29-35.
- [3] Hamada, T. (1993) A Bayesian sequential single machine scheduling problem to minimize the expected weighted sum of flowtimes of jobs with exponential processing times. *Operations Research* **41**, 924-934.
- [4] Naddef, D. and Santos, C. (1988) One-pass batching algorithms for the one-machine problem. *Discrete Applied Mathematics* **21**, 133-145.
- [5] Santos, C. and Magazine, M. (1985) Batching in single operation manufacturing systems. *Operations Research Letters* **4**, 99-103.
- [6] Shallcross, D. F. (1992) A polynomial algorithms for a one machine batching problem. *Operations Research Letters* **11**, 213-218.
- [7] 柳井 (2001) 学習を考慮した最適バッチスケジューリング. 星陵大論集 **33**, 331-341.