

有限資源の下でのファイルの最適バックアップ方策について

01503164 神戸商科大学 濱田年男 HAMADA Toshio

1 緒言

コンピュータやワードプロセッサを用いて種々のファイルを作成していく過程において、ファイルの修正と一定期間の保管という作業が繰り返されていく。この過程において、何らかの事故でファイルの内容を壊すことを避けるために、ファイルのバックアップファイルを作成する必要がある。しかしながら、バックアップファイルを作成することは面倒な作業であり、できるだけ省きたい作業である。このように、バックアップファイルを作成するか否かという問題は、作業期間を通して絶えずつきまとう問題である。このような問題、およびこれらに関連した研究は、Sandoh and Kawaii [1], Dohi, Aoki, Kaio and Osaki [2], Hamada [3] 等、その他にも多くの研究がある。

本研究では、利用できる資源としてのフロッピーディスクの利用可能枚数に制約がある場合に、バックアップファイルを作成する手間と、ファイルを壊す確率 p と、フロッピーディスクの利用可能枚数との関係を考えるためのモデルを構築して、動的計画法による解析を行い、最適戦略の構造を明らかにした。

2 モデル

マスターファイルが入った1枚のフロッピーディスクと、 n 枚の新しいフロッピーディスクが存在するものとする。マスターファイルの更新作業を一定期間毎に行っていくものとする。このとき、旧マスターファイルはそのまま保存し、新しいフロッピーディスクに更新されたマスターファイルを格納するものとする。このとき、新マスターファイルのバックアップコピーを作成しないという決定を a_0 、作成するという決定を a_1 とする。 a_0 を行うと、新しいフロッピーディスクの残り枚数は $n-1$ になり、また a_1 を行うと、新しいフロッピーディスクの残り枚数は $n-2$ になる。ここで、次の作業時点までにマスターファイルが壊れると、もしバックアップコピーが存在しなければ、旧マスターファイルにまで戻って作業を再開する必要が

あり、多大の労力を要するので、更新作業はここで終えることにする。しかし、もしマスターファイルのバックアップコピーが存在すれば、それを用いて作業を続行できるものとする。ただし、このバックアップコピーも壊れれば、そこで終了するものとする。これらの作業を続けていき、新しいフロッピーディスクがなくなれば作業はそこで終了するものとする。利用できる枚数が有限のフロッピーディスクを用いて、できるだけ長期間作業を進めるためには、バックアップコピーを作成するか否かの決定をどのように行えば良いかを決定する問題を考える。

3 定式化

新しいフロッピーディスクの残り枚数が n で、ファイルが壊れる確率を p とするとき、最適政策を用いたときの最大期待作業期間数を $F_n(p)$ とおくと、

$$F_0(p) = 0$$

$$F_1(p) = 1 - p$$

であり、 $n = 2, 3, 4, \dots$ に対して

$$F_n(p) = \{F_n^0(p), F_n^1(p)\}$$

が成立する。ここに $F_n^0(p)$ 、 $F_n^1(p)$ は、それぞれ新しいフロッピーディスクの残り枚数が n のときに、まず a_0 、 a_1 を行い、以後は最適政策にしたがったときの最大期待作業期間数であり、これらは

$$F_n^0(p) = (1-p)\{1 + F_{n-1}(p)\}$$

$$F_n^1(p) = (1-p^2)\{1 + F_{n-2}(p)\}$$

となる。今 $n = 2, 3, 4, \dots$ に対して

$$D_n(p) = F_n^0(p) - F_n^1(p)$$

と定義すると、 $D_n(p) \geq 0$ のとき、かつそのときのみ a_0 が最適となる。

補題 1. $0 < p < 1$ に対して

$$D_2(p) = (1-p)(1-2p),$$

$$D_3(p) = \begin{cases} (1-p)(1-4p+2p^2), & 0 < p \leq \frac{1}{2} \\ -p(1-p), & \frac{1}{2} < p < 1 \end{cases}$$

であり, また $n = 4, 5, 6, \dots$ に対して

$$D_n(p) = -p(1-p) + (1-p) \max\{D_{n-1}(p), 0\} \\ + (1-p^2) \min\{D_{n-2}(p), 0\}$$

が成立する.

証明. n についての帰納法による. \square

$F_n(p)$ と $D_n(p)$ に関する性質として, n に関する帰納法により, $n \geq 1$ に対しては,

$$\lim_{p \rightarrow 0^+} F_n(p) = n, \quad (1)$$

$$\lim_{p \rightarrow 1^-} F_n(p) = 0, \quad (2)$$

また $n \geq 2$ に対しては,

$$\lim_{p \rightarrow 0^+} D_n(p) = 1 \quad (3)$$

が成立する.

4 最適戦略の導出

p_n を以下のように定義する.

$$p_1 = 1$$

とし, $n \geq 2$ に対しては

$$p_n = 1 - \left(\frac{1}{2}\right)^{\frac{1}{n-1}}$$

と定義する. このとき, $p_1 \geq p_2 \geq p_3 \geq \dots$ が成立する. またこれらの値を計算すると,

$$p_2 = 0.5000$$

$$p_3 = 0.2929$$

$$p_4 = 0.2063$$

$$p_5 = 0.1591$$

$$p_6 = 0.1294$$

等が得られる.

これらの $p_n (n = 2, 3, 4, \dots)$ の値は, 最適政策において重要な役割を果たす. これについては, 次の定理が成立する.

定理 1. $n = 2, 3, 4, \dots$ に対して,
(i) $0 < p \leq p_{n-1}$ なる p に対して,

$$F_n^1(p) = \frac{1-p}{p} \{1 - (1-p)^n\}$$

および

$$F_n^2(p) = \frac{1-p^2}{p} \{1 - (1-p)^{n-1}\},$$

(ii) $0 < p \leq p_{n-1}$ なる p に対して

$$D_n(p) = 2(1-p) \left\{ (1-p)^{n-1} - \frac{1}{2} \right\},$$

(iii) $0 < p \leq p_{n-1}$ なる p に対して, $D_n(p)$ は p について狭義単調減少であり,

(iv) $0 < p \leq p_n$ のとき, かつそのときのみ $D_n(p) \geq 0$

(v) $0 < p \leq p_n$ のとき, かつそのときのみ $F_n(p) = F_n^1(p)$ である.

証明. n についての帰納法による. \square

この定理により, 最適政策は以下ようになる. 「残りのフロッピーディスクの枚数が n のとき, $0 < p \leq p_n$ ならば決定 a_0 を選択し, さもなければ決定 a_1 を選択するのが最適である.」

参考文献

- [1] Sandoh, H. and Kawai, H. (1991). An optimal N -job backup policy maximizing availability for a hard computer disk. *Journal of the Operations Research Society of Japan* 34, 383-390.
- [2] Dohi, T., Aoki, T., Kaio, N. and Osaki, S. (1997). Computational aspects of optimal checkpoint strategy in fault-tolerant database management. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* E80-A, 2006-2015.
- [3] Hamada, T. Optimal backup strategy for making files by computer. (Accepted).