パソコンによる豚閉鎖群育種集団における相加的血縁係数行列の逆行列算出のための記憶容量と演算時間に関する検討

佐藤正寛・小畑太郎

農林水産省畜産試験場, 筑波農林研究団地 305

(1995 年 2 月 15 日受付)

要 約 わが国で行われている豚の系統造成規模の大きさの集団に即して, 相加的血縁係数行列の逆行列 (A^{-1}) のノンゼロ要素だけを記憶させた場合とすべての要素を記憶させた場合の記憶容量と演算時間について検討した。まず, ノンゼロ要素数の理論値とその記憶容量 (ノンゼロ容量) を算出した。次に, 血縁情報をコンピュータシミュレーションによって発生させ, A^{-1} のパソコンによる算出時間について検討した。ノンゼロ容量は世代数や繁殖集団の大きさ等, データ量に比例する傾向にあり, A^{-1} のすべての要素を記憶させた容量に比べて非常に小さかった。 A^{-1} のノンゼロ要素だけを記憶させた場合の演算時間は, A^{-1} の要素をすべて記憶させた場合のそれに比べてやや短い傾向にあった。

緒 言

近年, 種畜の遺伝的能力 (育種価) を評価する方法として, アニマルモデルによる最良線形不偏予測 (BLUP) 法¹⁾ が広く用いられるようになってきた。一般に, この方法による育種価の BLUP の算出にはキングサイズの方程式を解く必要があるため, 大型コンピュータやワークステーションが用いられている²⁾。しかし, 豚の系統造成を行っている場所や種豚農家において種豚を評価しようとする場合, ワークステーション等のコンピュータの利用は困難であることが多い。

コンピュータによって育種価の BLUP を算出する場合, 相加的血縁係数行列の逆行列 (A^{-1}) を作成するときおよび BLUP 方程式の解を得るときに最も多くの記憶容量と演算時間を必要とする。このとき A^{-1} の要素のうち, ゼロでない要素 (ノンゼロ要素) が少なければ, ノンゼロ要素とその位置を示す間接番地だけを記憶させることによって, A^{-1} の記憶容量を大幅に減少させることが可能である^{3,4)}。

本研究では, わが国で行われている豚の系統造成規模

の集団において, A^{-1} をパソコンによって求めるための記憶容量および演算時間について検討した。

方 法

わが国で行われている豚の系統造成規模の繁殖集団の大きさとして, ①雄 5 頭雌 25 頭, ②雄 10 頭雌 50 頭, ③雄 15 頭雌 75 頭の 3 通りを設定した。一腹あたりの育成頭数はそれぞれの規模ごとに 3~7 頭とし, 一腹あたりの育成頭数が偶数のときは雌雄同数を, それが奇数のときは雌が雄よりも 1 頭多くなるように育成することとした。世代数は 2~10 とした。

まず, HENDERSON⁵⁾ のアルゴリズムによって, A^{-1} のノンゼロ要素数の理論値を求めた。このとき, A^{-1} の要素 (1 数値倍精度=8 バイト) をすべて記憶させた場合 (a) とノンゼロ要素およびその間接番地を記憶させた場合 (b) の記憶容量を算出した。なお, ノンゼロ要素の間接番地は, 第 i 行のノンゼロ要素数および第 i 行のノンゼロ要素の列で表すこととした。したがって, 2 種類の間接番地のうち, 1 種類の数値は A^{-1} の行数, 他の 1 種類のそれはノンゼロ要素の数 (それぞれ 1 数値=4 バイ

A Study of Memory Size and Computation Time of a Personal Computer for Inverse of an Additive Genetic Relationship Matrix in a Closed Herd of Swine

M. SATOH and T. OBATA

National Institute of Animal Industry, Tsukuba Norinkenkyudanchi, 305

ト)が必要である。ただし、 A^{-1} は対称行列であるため、 A^{-1} の上三角行列の要素のみを記憶させた。

次に、繁殖集団の大きさを雄 10 頭、雌 50 頭、世代数を 5 および 10 とし、一腹あたりの育成頭数を 3~7 頭とした 10 通りの血縁情報をコンピュータシミュレーションによって発生させ、 A^{-1} の要素をすべて記憶させた場合 (a) とノンゼロ要素およびその間接番地を記憶させた場合 (b) の演算時間と記憶容量の比較を行った。シミュレーションの条件は、

- (1) 雌親の分娩率~N (0.90, 0.05²),
- (2) 選抜は無作為,

とした。 A^{-1} は HENDERSON⁵⁾ と QUAAS⁶⁾ のアルゴリズムを併用して求めた。すなわち、 A^{-1} を三角行列分解すると、

$$A = (T')^{-1} D^{-1} T^{-1}$$

が一意に定まる。ここで、 T^{-1} は上三角行列、 D^{-1} は対角行列、'は行列の転置を表す。この D^{-1} の要素を HENDERSON⁵⁾ の方法によって求めた後、QUAAS⁶⁾ の方法によって A^{-1} を算出した。

シミュレーションと A^{-1} の算出には、NEC のパーソナルコンピュータ PC-9821 Ap 2 (CPU=i 486) を用いた。

結 果

A^{-1} の要素をすべて記憶させた場合 (a) とノンゼロ要素だけを記憶させた場合 (b) の記憶容量の比較を表 1 に示した。記憶容量は腹あたりの育成頭数にほぼ比例したため、表 1 には腹あたりの育成頭数が 4 頭の場合を示

している。(a) は世代数の増加に伴って急激に増加し、特に繁殖集団が大きいくほど (a) の増加量が大きくなった。(b) は世代数や繁殖集団の大きさに比例する傾向にあった。また、その大きさは (a) に比べて非常に小さく、繁殖集団の大きさが雄 15 雌 75、腹あたりの育成頭数が 4 で 10 世代分 (2,790 頭) の A^{-1} を記憶させるためには、(a) で 30 M バイトあまりのメモリが必要であるのに対し、(b) では 0.12 M バイトのメモリで記憶させることが可能であった。表 1 には示していないが、(b) では腹あたりの育成頭数が 7 のときでも、0.20 M バイトのメモリがあれば A^{-1} を記憶させることが可能であった。

A^{-1} の要素をすべて記憶させた場合 (a) とノンゼロ要素だけを記憶させた場合 (b) のシミュレーションによる記憶容量と演算時間の比較を表 2 に示した。このときの最大個体数は 3,012 頭であった。理論計算でも予想された通り、(b) は (a) に比べて記憶容量を大幅に節約することができた。一方、演算時間は個体数が多くなるにつれて急激に長くなった。また、(b) は (a) に比べて演算時間が短くなるものの、その差は極めて小さかった。

考 察

近年のコンピュータの発展は目ざましく、パソコンにおいてもハードディスク等の大きな記憶容量を搭載したものが市販されている。そこでアニマルモデルによる育種価の予測をパソコンによって行うことができれば、系統造成関係場所はもちろん、農家段階においても育種価の予測が可能となり、豚の育種効率率は飛躍的に向上する

表 1. 相加的血縁係数行列の逆行列 (A^{-1}) の要素をすべて記憶された場合 (a) とノンゼロ要素だけを記憶させた場合 (b) の記憶容量 (メガバイト) の理論値の比較

Table 1. Comparison of theoretical memory sizes (mega-byte) between the storage of all elements (a) and of non-zero elements (b) of the inverses of the additive relationship matrices (A^{-1}).

No. of generation	5 sires and 25 dams			10 sires and 50 dams			15 sires and 75 dams		
	No. of animals	(a)	(b)	No. of animals	(a)	(b)	No. of animals	(a)	(b)
2	130	0.07	0.01	260	0.26	0.01	390	0.58	0.02
3	230	0.20	0.01	460	0.81	0.02	690	1.82	0.03
4	330	0.42	0.02	660	1.66	0.03	990	3.74	0.05
5	430	0.71	0.02	860	2.82	0.04	1,290	6.35	0.06
6	530	1.00	0.02	1,060	4.29	0.05	1,590	9.65	0.07
7	630	1.52	0.03	1,260	6.06	0.05	1,890	13.63	0.08
8	730	2.04	0.04	1,460	8.14	0.06	2,190	18.30	0.09
9	830	2.63	0.04	1,660	10.52	0.07	2,490	23.66	0.11
10	930	3.30	0.04	1,860	13.20	0.08	2,790	29.70	0.12

The number raising pigs per litter was assigned as 4.

表 2. 相加的血縁係数行列の逆行列 (A^{-1}) の要素をすべて記憶させた場合 (a) とノンゼロ要素だけを記憶させた場合 (b) の記憶容量 (MB) と演算時間 (秒) のシミュレーションによる比較

Table 2. Comparison of memory sizes (mega-byte) and operation time (second) between the storage of all elements (a) and of non-zero elements (b) of the inverses of the additive relationship matrices (A^{-1}) by computer simulations.

No. of generation	Raizing pigs per litter	No. of animals	(a)		(b)	
			Memory size (MB)	Operation time (sec.)	Memory size (MB)	Operation time (sec.)
5	3	619	1.49	50	0.03	50
	4	806	2.52	106	0.03	102
	5	993	3.82	192	0.04	182
	6	1,180	5.40	318	0.05	304
	7	1,367	7.25	482	0.06	462
10	3	1,324	6.80	432	0.06	424
	4	1,746	11.82	984	0.07	944
	5	2,168	18.22	1,844	0.09	1,770
	6	2,590	26.00	3,108	0.11	3,002
	7	3,012	35.16	4,850	0.12	4,688

ことが期待される。

本研究の結果から、ノンゼロ要素だけを記憶させることにより、3,000 頭規模の A^{-1} の要素でも 0.1 M バイト程度のメモリがあれば記憶可能なことが明らかとなった。また、このときの演算時間も 80 分前後であったことから、わが国で行われている豚の系統造成規模の集団であれば、パソコンによって A^{-1} を算出し、記憶させることは十分可能であることが明らかとなった。

単形質の BLUP 方程式の係数行列のうち変量効果に相当する部分は、 I を単位行列、 k を任意の定数とすると、一般に $I+kA^{-1}$ として表すことができ、この部分のノンゼロ要素数は A^{-1} のノンゼロ要素数に等しい。一方、豚の閉鎖群育種では環境要因が比較的斉一であるため、一般に母数効果として取り上げる要因あるいはその水準数は少ないと考えてよい。したがって、変量効果に相当する部分のノンゼロ要素は、BLUP 方程式の係数行列のノンゼロ要素の大部分を占めることになる。このことは A^{-1} を記憶させることが十分可能であれば、BLUP 方程式の係数行列を保存することも容易であることを示している。

多形質のアニマルモデルでは、本研究で算出した記憶容量におおよそ形質数の二乗を乗じた記憶容量が必要になる。例えば、5 形質のアニマルモデルでは、本研究で算出した容量の約 25 倍の記憶容量が必要になる。この場合、閉鎖集団では世代数を 3~5 世代程度以上さかのぼっても選抜の正確度は大きく変わらないこと⁷⁾ から、

初期世代の個体を除くことによって個体数を縮小すれば記憶容量を節約することができる。また、5 形質 3,000 頭規模でも必要なメモリは 3 M バイト程度であることから、ハードディスク等の補助記憶装置を用いれば、多形質のアニマルモデルでも BLUP 方程式の係数行列を記憶させることは容易である。さらに、育種価の予測に用いる数学モデルによっては、正準変換を行うことによって BLUP 方程式のサイズを小さくすることが可能である⁸⁾。以上のことから、パソコンを用い、多形質のアニマルモデルで種畜評価を行うことは、記憶容量の観点からは十分可能であると考えられる。

今後はパソコンによって BLUP 方程式の解を得るための演算時間について検討する必要がある。

文 献

- 1) HENDERSON, C.R.: Anim. Breed. Genet. Sym. in Honor of Dr. LUSH, 10-41 A.S.A.S. and A.D.S.A., Illinois, 1973.
- 2) LONG, T., H. BRANDT and K. HAMMOND: Pig News and Information, 12, 217-219, 1991.
- 3) DUFF, I.S., A.M. ERISMAN and J.K. REID: Direct methods for sparse matrices, 19-39, Clarendon Press, Oxford, 1989.
- 4) DONGARRA, J.J., I.S. DUFF, D.C. SORESENSEN and H. A. van der VORST: 小国力訳, 連立一次方程式の解法, 97-127, 丸善, 東京, 1993.
- 5) HENDERSON, C.R.: J. Dairy. Sci., 58, 1727-1730, 1975.

- 6) QUAAS, R.L. : Biometrics, **32**, 949-953, 1976.
7) 佐藤正寛・西田 朗 : 日畜会報, **61**, 902-906, 1990.
8) ARNASON, T. : Acta Agric. Scand., **32**, 171-176, 1982.

**A Study of Memory Size and Computation Time of a Personal
Computer for Inverse of an Additive Genetic Relationship
Matrix in a Closed Herd of Swine**

Masahiro SATOH and Taro OBATA

National Institute of Animal Industry, Tsukuba Norinkenkyudanchi, 305

We compared memory size and computation time required for the storage of non-zero elements (SNZE) and all elements (SAE) of the inverses of the additive relationship matrices ($\mathbf{A}^{-1}$) assuming the standard population size of the closed herd of swine breeding program in Japan. The number of non-zero elements and their memory sizes of the $\mathbf{A}^{-1}$ were theoretically calculated, and the computation time on SNZE and SAE of the $\mathbf{A}^{-1}$ of pedigree data which were simulated by computer was investigated. The memory size of SNZE linearly increased with increasing amount of data, such as the number of generation and population size. The memory size of SNZE was much smaller than that of SAE. The computation time for SNZE was slightly shorter than that for SAE.

Jpn. J. Swine Science, **32**, 3 : 106-109

Key words : memory size, computation time, personal computer, additive matrix, closed herd