

固有値計算と連立一次方程式のプログラムを用いた コンピュータの性能評価

香田 美紀、長嶋 雲兵

お茶の水女子大学 理学部 情報科学科

科学技術計算でよく用いられるLU分解法を用いた連立一次方程式の計算と、QR法を用いた固有値問題のプログラムを用いて、HP710,IBM580,HITAC 3050Rのワークステーションと、CRAY C30,HITAC S3800,NEC SX-3R, TITAN 2-800, CONVEX C3840のスーパーコンピュータでCPU時間を測定し、それぞれの性能を比較した。二種類の違う問題（連立一次方程式と固有値計算）を扱った場合、カタログ性能からは予測できない程の性能比の違いが見られ、行列の種類の違いによる性能比の違いも見られた。また、扱う行列のサイズを変えたり、行列の種類を変えてみることで、性能比が違うという結果も得られた。

Performance evaluation of Workstation and Supercomputer --- Eigenvalue problem and Linear equation ---

Miki Kouda and Umpei Nagashima

Department of Information Sciences Faculty of Science Ochanomizu University

In order to study the CPU performance of Workstations and Supercomputers: (HP710,IBM580 and HITAC 3050R, and Supercomputer : Cray C90(1 cpu),HITAC S3800(1 cpu), NECSX-3R (1 cpu), TITAN2-800(1 cpu), CONVEX C3840(1 cpu)), CPU time has been measured using newly coded programs for Eigenvalues problem and Linear equation. For Workstation, the size of matrix is varied in the range from 50 to 250, which interval is 50. That for Supercomputer is from 100 to 1000, which interval 100.

For different problem, the ratio of CPU performance is different from that estimated the performance on catalog. In the case of large size matrix, strong size dependency has not been observed on Supercomputer.

1 はじめに

1.1 ベンチマークテスト

ベンチマークの目的は、實際上システム設計の立場からのゴールとしてまた、計算機を用いた問題解決を図る立場から見た、スタート地点としての二面性がある。問題解決という立場から言えば、プログラムを通していくつかのシステムで問題を解くことにより、それぞれのシステムの解かれる問題に対する適合性の問題が重要となり、さらに、ベンチマークテストの結果から、新しい計算モデルに対応した問題解決方法を考えることができる。また、システムを設計する立場から言えば、いくつかの異なるプログラムを通して、いくつかの鍵となる部分の性能を見ることにより、さまざまなプログラムに対するシステムアーキテクチャの適合性についての知見を得ることができる。さらに、その結果から新しい計算モデルに基づいた計算技法に対応するためにシステムのどこを改良するべきかを知ることができる。

一般にベンチマークテスト自体はある種の指標を与えるにすぎないが、実際の応用プログラムに実行された数値が結果として示される事は少なく、たとえ実際の応用プログラムをベンチマークテストに適用したとしても、それは、ベンチマークテストという性格上極めて、小規模のものにならざるを得ない。実際には、大規模な問題に計算機が適用されるときも、ベンチマークのスケラビリティの確認なく調査が行われているのが現状である。一般には、ベンチマークテストにおけるスケラビリティは保証されていると信じられており、確かに従来の逐次型計算機ではスケラビリティの保証があったが、現在では、パイプラインや先行制御による命令の並列実行が普通であり、スケラビリティの保証は定かではない。残念ながら、その調査の例は極めて少ない。

本論文では、アプリケーションの立場、すなわち計算機を用いて問題解決を図るという立場から、実際の問題に多く使われているLU分解法による連立一次方程式と、QR法による固有値問題を解くプログラムを用いて、いくつかのワークステーションとスーパーコンピュータの個性を明らかにするために、問題のサイズを変えながらその性能の変化を見た。

1.2 評価指標

従来よりベンチマークプログラム実行時間はプログラムと入力データに固有の値であり、一般性に乏しいので、むしろ実行時間の値とプログラムの実行命令数や実行演算数などを用いて、ある種のモデルに基づく評価指標を求め、それを用いて計算機の性能が議論されることが多い。なかでも、現在広く用いられている評価指標として次のようなものがある。

- MFLOPS (Million Floating-point Operations Per Second)
1秒間あたりの浮動小数点演算回数を次式のように 10^6 単位で表す。
$$\text{MFLOPS} = (\text{浮動小数点演算数}) \div \{(\text{プログラム実行時間}) \times (10^6)\}$$
- SPECint92
整数演算用の6本のCプログラム群で構成されたCINT92における結果の幾何平均値。
- SPECfp92
12本がFORTRAN、2本がCで書かれた、9本の倍精度浮動小数点演算プログラム群で構成されたCFP92における結果の幾何平均値。
- LINPACK 値 (MFLOPS 倍精度)
 100×100 の線形方程式の結果。
- ベクトルアーキテクチャの評価指標
演算装置の処理性能 (スループット, 数), r_{∞}
メモリから演算装置への転送能力, $n_{1/2}$

ベンチマークを行ったワークステーションとスーパーコンピュータのカタログ上のこれらの評価指標の値は表1、表2に示す通りである。ただし、HITAC 3050Rの値は、我々の持っているテストプログラムの結果であり、メーカー保証の値ではない。各種の計算機について、通常これらの値はカタログに示されている。しかしながら、一般に大規模な計算を行うようなとき、どのような問題を扱うか、また、どれくらいの規模の計算を行うかによって実際の計算時間とカタログ性能から予測される時間と違いがみられることは、広く知られている。にもかかわらず、一般の応用プログラムを実行させるための計算機の性能は、それらの指標のみを用いて議論されることが多い。

Computer name	SPECint92	SPECfp92	MFLOPS
HP710	31.6	47.6	12.2
IBM580	61.9	134.6	38.1
HITAC3050R	31.6	47.6	12.2

表 1: 測定に用いたワークステーションのカタログ性能値

Computer name	Linpack MFLOPS	Best Effort MFLOPS	Theoretical Peak MFLOPS
CONVEX C3840	75	—	120
TITAN2-800	10	25	32
HITAC S3800/480	—	—	8000
NEC SX-3/34R	—	—	6400
CRAY C90	387	902	952

表 2: 測定に用いたスーパーコンピュータの Linpack 性能値

実際の問題を解く時間の性能を実際の応用プログラムを用いて示すために、本研究では実際の問題を解くことで得られる CPU 時間からこれらの評価指標を計算することをせずに CPU 時間そのものを見ることにした。なぜならば、これらの指標は、ある種の計算モデルに基づいており、これを計算することは、そのモデルに含まれる不確定要素をも評価の中にも含み込むことになるからである。直接的に特定の各問題に対する各計算機の性能を見ることが本研究の第一の目的である。

また、一般に科学技術計算においては、何時間もかかるような計算を取り扱う場合が多いので、計算実行時間の事前予測は、実際に計算を実行する上で非常に大きな意味を持つ。そこで、問題の大きさを変えた時の性能の変化を見ることも本研究の目的である。

本研究では IBM580, HP710, HITAC 3050R という三種類のワークステーションと、NEC SX-3/34R, HITAC S3800, CRAY C90, TITAN 2-800, CONVEX C3840 という 5 種類のスーパーコンピュータを用い、実際の応用に良く用いられる連立一次方程式と固有値問題のプログラムを用いて、問題の違いや大きさの違いによる性能変化を見る調査を行った。

2 測定方法

2.1 プログラムの概要

実際に計算機上で処理される問題の多くは、行列演算であり、なかでも固有値問題と連立一次方程式が数多く

解かれている。そこで本研究では、ハウスホルダーQR法を用いて固有値・固有ベクトルを求めたときの CPU 時間と、解ベクトルの初期値がすべて 1 として LU 分解法を用いて連立一次方程式を解いた場合の CPU 時間を測定した。

用いた行列は、フランク行列、ブロック行列、ランダム行列、ヒュッケル行列である。それぞれの行列要素の定義は以下の通りである。

フランク行列

$$F(i, j) = N + 1 - \max(i, j)$$

ブロック行列

$$B(i, j) = 4\delta(i, j) - \sum_{k=1}^{r-1} \sum_{j=0}^{r-2} [B]$$

$$\text{where } r = \lceil \sqrt{N} \rceil$$

$$[B] = \delta(i, k+1+r)\delta(j, i+r-1) + \delta(i, k+r)\delta(j, i+r+1) + \delta(i, k+1+r)\delta(i, j+r-1) + \delta(i, k+r)\delta(i, j+r+1)$$

ランダム行列

$$-32768 < R(i, j) < 32768$$

ヒュッケル行列

$$H(i, j) = -7.2\delta(i, j) - 3.0/(i-j)^2(1 - \delta(i, j)) \\ (i = j \text{ のとき } -7.2)$$

計算した行列の大きさは、ワークステーションにおいては 50 から 250 を 50 おきにとった 5 点で、スーパーコンピュータにおいては 100 から 1000 を 100 おきにとった 10 点である。計算精度は、 10^{-16} である。

用いたベンチマークプログラムは、フォートランを用いて今回新たに作成した。プログラムのコンパイルは、最適化オプションを用いて行った。時間測定は、secsnd 関数（機種によっては、clock, second, mclock ルーチン等を用いた）をプログラム内部から呼び出すことにより、CPU 時間または Wallclock 時間を求めた。

2.2 用いたコンピュータについて

測定に用いたワークステーションは、お茶の水女子大学情報科学科の HP710 と IBM580 , および岡崎国立共同研究機構分子科学研究所電子計算機センターの HITAC 3050R の三種の計算機である。それぞれ 64 Mbyte の主記憶装置、2.3 Gbyte のディスク装置を持っている。

スーパーコンピュータは、東京エレクトロンの CONVEX C3840, クボタコンピュータの TITAN 2-800, 東京大学計算機センターの HITAC S3800, 岡崎国立共同研究機構分子科学研究所電子計算機センターの NEC SX-3R, CRAY リサーチパーク (ミネアポリス) の CRAY C90 の五種の計算機である。それぞれ 1 CPU の性能を調べた。CONVEX C3840 は 512 Mbyte, TITAN 2-800 は 128 Mbyte, NEC SX-3R は 2 Gbyte, HITAC S3800 は 2048 Mbyte, CRAY C90 は 512 Mbyte の主記憶装置を持っている。

3 結果と考察

3.1 ワークステーションにおける結果

得られた結果は、膨大なものであり、紙面の都合上ここからは、フランク行列による結果をあげていく。

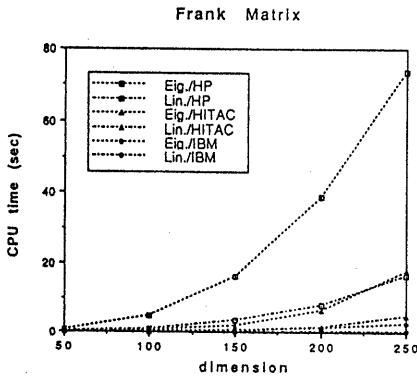


図 1: 固有値計算と連立一次方程式にかかった CPU 時間

図 1 は、ハウスホルダー QR 法を用いてランダム行列の全固有値・固有ベクトルを求めたとき (Eig.)、およ

び解ベクトルの初期値がすべて 1 として LU 分解法を用いて連立一次方程式を解いた場合 (Lin.) の実測された CPU 時間を次元数に対して示したものである。

ほぼ次元数の 3 乗で計算時間が増加している。いずれのマシンも、連立一次方程式よりも固有値計算の方が CPU 時間は大きくなっている。しかし、HITAC 3050R では、次元数が増えたときに他よりも時間の増え方が大きくなっている。

図 2 は、行列要素の違いにおける CPU 時間の変化を見るために、HP710 で固有値計算を行なった時の、フランク行列とランダム行列の場合を示したものである。

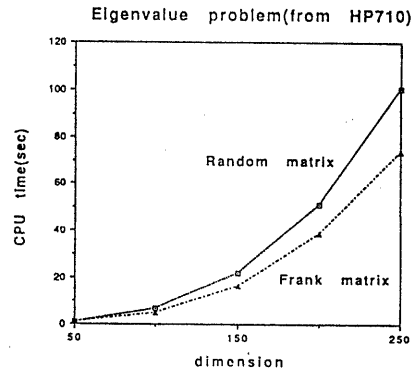


図 2: フランク行列とランダム行列における固有値計算の CPU 時間 (HP710)

ランダム行列の方が行列要素に規則性がなく、また、対角要素と比対角要素の絶対値の差も小さいことから、行列の計算としては非常に難しい問題となっていて CPU 時間が大きい。

図 3 は、IBM580 と HP710 と HITAC 3050R での連立一次方程式の問題と固有値問題の CPU 時間の比 (固有値問題/連立一次方程式) を示したものである。

どのマシンも、次元数が大きいところではほぼ一定の比をとっている。つまり、次元数が大きくなると両者では行列の次元数が増えてもそれぞれの問題にかかる時間の増加の割合は変わらないということである。これは、次元数が増えた時に CPU 時間は同じオーダーで増加している。この問題に関しては、次元数が大きくなると、ほぼスケラビリティが満たされている。また、IBM 580 のグラフが他のグラフよりほぼ上にあることから、IBM 580 は、固有値問題より連立一次方程式のほうを

得意としている事が分かる。HITAC 3050R は、次元数が増えるにつれて多少右下がりになっている。先の図1にも示されているように、HITAC 3050R の場合次元数が増えるとかかる時間が急激に増えていることを反映している。図3からは、HITAC 3050R は問題の規模が大きくなると連立一次方程式を解く場合のほうが性能の低下が著しくなることが読みとれる。詳細に分析する手段がないので確かなことは言えないが、たぶんこれは、HITAC 3050R のメモリ管理アーキテクチャに問題があるのではないだろうか。少なくとも相対的な性能を見る限り、HITAC 3050R は、この領域で、スケラビリティが保証されていない。

次元数が小さな所では、複雑な現象が見られることが予想されるが、よりキザミの小さな測定がなされなかったので、ここでは議論しない。

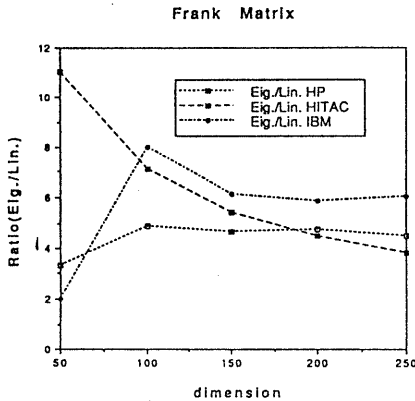


図3: 固有値計算と連立一次方程式のCPU時間の比

図4は、IBM580を1としてHP710とHITAC 3050Rのそれぞれの問題に対するCPU時間の比を示したものである。

まず、HP710とIBM580で比較してみる。IBM580はHP710に比べカタログ性能のLINPACK MFLOPSで3.1倍速く、SPECfp92では2.8倍速いことになる。他方実測のでは、4つのグラフからみて分かるように、両者の性能比は固有値問題では、4つの行列での結果を見ると、25~30倍となり、カタログ性能比の10倍ほどの性能がある。この違いは、IBM580での最適化がかなり利いているためと思われる。それらは、次元数の変化に対してもほぼ一定の値を示している。この両者では、従来のMFLOPSやSPECfpといっ

たカタログ性能は、正しい性能比を示しておらず、固有値問題に関してほぼ実際の性能比の10倍以上の性能比がある。

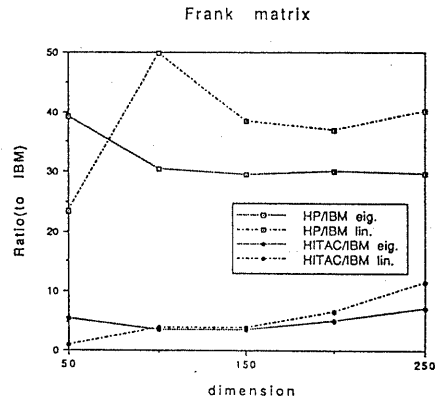


図4: IBM580に対するCPU時間の比

連立一次方程式の問題では、実測の場合、両者の性能比は40~50倍となっているが、次元数が増えるに従って比が一定になってくる。また、連立一次方程式問題の性能比の方が固有値問題に比べて大きいことからIBM580はHP710に比べて、連立一次方程式の問題が得意なことが分かる。これらの事柄は、カタログ性能の比較からは全くうかがいしれない事柄である。

IBM580とHITAC 3050Rでは、固有値計算と連立一次方程式ではほぼ同じ性能比をとっている。性能比は実測で4~5倍である。これは、ほぼカタログ性能の比に近い値である。また、次元数の変化に対してもほぼ一定の値を示している。次元数の増加に伴って、連立一次方程式の方が上になっていることから、次元数が大きくなると連立一次方程式を不得意としてくることがわかる。

この結果から、HP710と、HITAC 3050Rの性能を比較してみると、ハードウェア性能は、ほぼ同じであるが、実行性能は、HITAC 3050Rの方がHP710より速い。これは、コンパイラなどの基礎ソフトの性能の違いが現れているものと想像できる。

これらの結果は、従来の性能指標に基づくカタログ性能からは全く予測できないものである。

3.2 スーパーコンピュータにおける結果

図5は、ハウスホルダーQR法を用いて全固有値・固有ベクトルを求めたときおよび、解ベクトルの初期値がすべて1としてLU分解法を用いて連立一次方程式を解いた場合のCPU時間を次元数に対して示したものである。5種類のスーパーコンピュータで測定しているが、かかった時間にかかなりの差があったため、この図ではHITAC, NEC, CRAYの3種の場合をあげている。

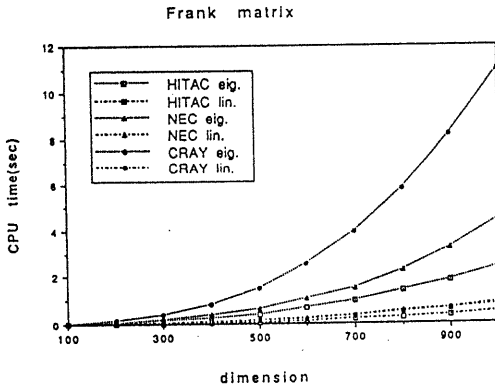


図5: 固有値計算と連立一次方程式にかかったCPU時間

それぞれの場合、ほぼ次元数の3乗で計算時間が增加されている。またどのグラフからも、ワークステーションの時と同様、連立一次方程式より固有値問題の方がCPU時間がかかっている。

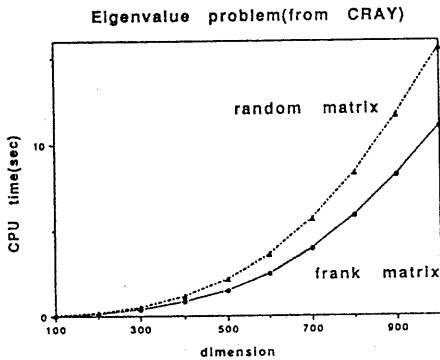


図6: フランク行列とランダム行列における固有値計算のCPU時間(CRAY)

次に図6は、行列要素の違いにおけるCPU時間の変化を見るために、CRAYで固有値計算を行なった時の、フランク行列とランダム行列を比較した場合である。

これもワークステーションの時と同様、ランダム行列の方が行列要素に規則性がなく、また、対角要素と比対角要素の絶対値の差も小さいことから、行列の計算としては非常に難しい問題となっていてCPU時間が大きい。

次に図7に連立一次方程式の問題と固有値問題のCPU時間の比(固有値問題/連立一次方程式)を示す。

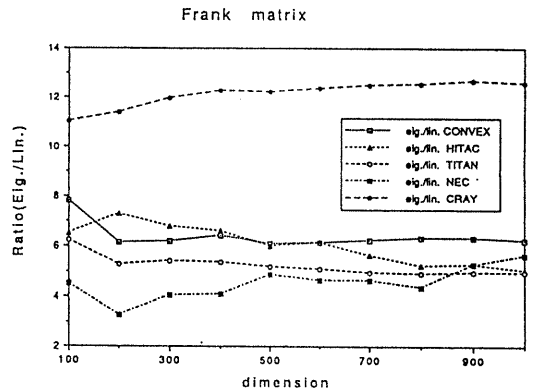


図7: 固有値計算と連立一次方程式のCPU時間の比

まず、CRAYのグラフが他よりも上にあることから、CRAYは連立一次方程式を得意とすることがわかる。また、HITAC, NECは多少ばらつきがあるが、次元数の大きいところで、どの計算機でも2つの問題間の演算量の比はほぼ一定になっている。

図8, 9は、HITACを1としてその他のコンピュータのそれぞれの問題に対するCPU時間の比を示したものである。これも、比の値にかかなりの差が見られたため分けてグラフを作成した。

TITANは、表2における最大ベクトル性能比では250倍遅い性能になるが、次元数に比例するように10-90倍遅いだけとなり、HITACに対する相対性能は良くなっている。HITACに比べて次元数の増加に対して性能の低下が著しくなっているが、カタログ性能よりは良い性能を出している。また、連立一次方程式のグラフの方が上にあることから、固有値計算の方をHITACに比べて得意としている。

表2に見られるように CONVEX の、HITAC との最大ベクトル性能比では67倍になるが、TITANほど相対性能の増加はないものの、10-30倍と増加している。やはり HITAC に比べれば性能の低下は見られる。しかし、固有値計算と連立一次方程式では、ほとんどグラフに差は見られない。次元数が600のところまで2つのグラフが交差していることから、次元数が大きくなると HITAC に比べて連立一次方程式の計算を得意とするようになることがわかる。

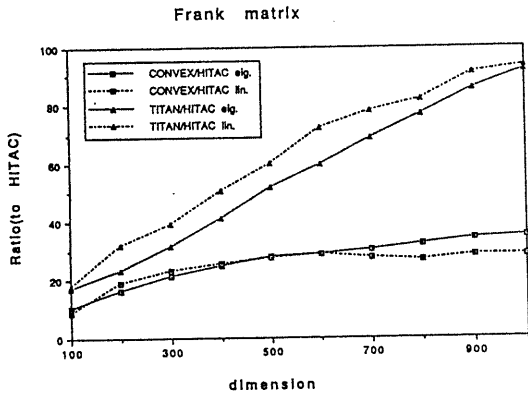


図 8: HITAC に対する CPU 時間の比 (CONVEX , TITAN)

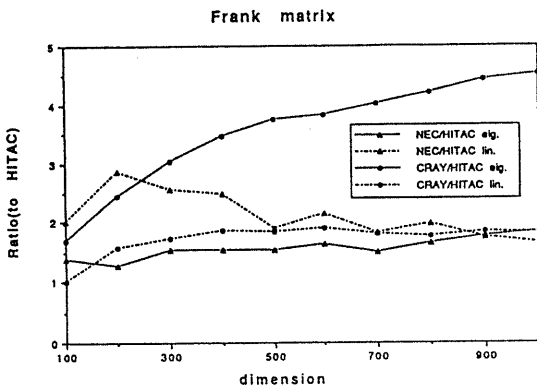


図 9: HITAC に対する CPU 時間の比 (NEC , CRAY)

次に図9で、NECを見ると、表2における最大ベ

クトル性能比では1.3倍になり、実測では連立一次方程式で約2倍、固有値計算で約1.5倍の性能比が見られ、ほぼカタログ性能から予測される相対性能通りになっている。

CRAYを見ると、表2における最大ベクトル性能比では8.4倍になるが、固有値計算と連立一次方程式のグラフでかなりの差が見られる。連立一次方程式では、約1.5倍で一定の比をとっているが、固有値計算では、次元数が増えるにつれて1.5倍から4倍まで増加している。これから、HITACに比べて固有値計算の方が不得意で、特に次元数が増加すると相対性能が低下することがわかる。

ここで取り上げている計算機はベクトルパラレル機であるので、ベクトル計算機の性能指標として有名な r_{∞} と $n_{1/2}$ でそれぞれ性能を議論してみる。

LU分解で連立一次方程式を解く場合の演算回数は、

$$N_{flop} = \frac{2}{3}N^3 + O(N^2)$$

であるので、本研究で得られた実測値を用いて、評価指標であげた r_{∞} と $n_{1/2}$ を見積もることができる。CPU時間を3次の多項式で最小自乗近似すると、

$$t = c_3N^3 + c_2N^2 + c_1N + c_0$$

の係数 c_3, c_2, c_1, c_0 を得ることができる。この係数を用いると、

$$r_{\infty} = (2/3)c_3$$

$$n_{1/2} = c_2/c_3$$

となる。これらの実測から得られる性能比を表3にあげる。

Computer name	Frank	
	r_{∞}	$n_{1/2}$
CONVEX C3840	23.7	0.012
TITAN2-800	102.0	0.566
NEC SX-3/34R	0.022	230.3
CRAY C90	1.37	0.37
HITAC S3800/480	1.0	1.0

表 3: $r_{\infty}, n_{1/2}$ の値

図8,9の比と比較すると、これらの実測から得られる性能比は CONVEX は、 r_{∞} が固有値計算、連立一次方程式ともに良い見積りを与えている。TITAN は、

比の値の増加が著しいので、この値は参考にできない。CRAY は、 r_{∞} が連立一次方程式の見積りを与えている。これ以外は、特に良い見積りは与えていない。

4 結論

本論文では、LU分解を使った連立一次方程式の計算と、ハウスホルダーQR法を使った固有値問題をベンチマークテストとして用いた。本論文で示したようにワークステーション、スーパーコンピュータ共に、それぞれ問題によって得意不得意がある。

今回用いたワークステーションについてまとめると次のようなことがいえる。問題が違った場合 IBM 580 と HP 710 は得意不得意はあるにしても、次元数の変化に対して CPU 時間の比はほぼ一定に保たれる。ところが HITAC 3050R は、カタログ性能から期待される速度より速い性能を示すことがあるが、問題やサイズの違いでかなり性能にばらつきがある。

今回用いたスーパーコンピュータについてまとめると次のようなことがいえる。問題による得意不得意の差が一番大きかったのは、CRAY C90 である。CONVEX C3840 と TITAN 2-800 は、他の HITAC S3800 , NEC SX-3 , CRAY C90 に比べ、計算にかかった CPU 時間にかかなりの差があったため細かいところまで比較することは困難であった。また、NEC SX-3 以外は問題が違った場合にでも次元数の変化に対して CPU 時間の比はほぼ一定に保たれる。NEC SX-3 は、問題やサイズの違いでかなりばらつきがみられる。特に、ブロック行列とヒュッケル行列で、ばらつきが大きく、これは計算時の収束性等の問題もあると思われる。

HITAC S3800 との比から、NEC SX-3 は多少のばらつきはあるものの、性能比がほぼ一定であるが、CRAY C90 は、問題によってかなりの違いが表れる。また、CONVEX C3840 , TITAN 2-800 は、問題による差、次元数による差はかなり大きい。

また、ワークステーションとスーパーコンピュータ共に、当然のことながら、行列要素に規則性がないものや行列が複雑なものの方がより時間がかかる。

このように、カタログ性能値から予測できるものと実測での結果とはかなりの差が見られ、良く設計されたベンチマークテストはこのような予測を行なうためには必要不可欠である。

5 おわりに

問題の種類や大きさの違いによって計算機の性能比は変わっている。しかし、ここで用いた問題やプログラムはほんの一例にすぎず、プログラムの作り方などによってまた違った結果が得られることもある。この様な限られたベンチマークで得られた値からだけでは計算機システムの性能は明らかにできないが、性能評価を行う時に、どのようなプログラムを用いどのようなデータを用いるかによっては、性能が極端に良くなったり悪くなったりということがあることがわかる。

自分の持つ問題に対する計算機の性能をよりよく知るためには、自分の行いたい問題の計算の内容が、どのようなベンチマークテストと相似であるかを知ることが重要である。そのため、今後の課題としては、いくつかの典型的なベンチマークテスト間の相似の調査方法及び、いくつかのスーパーコンピュータの性能について研究を深めていく必要がある。

謝辞

本研究をまとめるにあたり、お茶の水女子大学理学部情報科学科細矢治夫教授、および論文作成に至るまであらゆる面で御指導、御助言を頂きました、長嶋雲兵助教授には大変感謝致します。また、本研究に御協力頂きました、東京エレクトロンの只野勇氏、雪竹潤氏、日本クレイの福島 信弘氏、日本IBMの後藤隆氏、クボタコンピュータの新国竜也氏、矢地田工氏、分子科学研究所電子計算機センターの皆様、東京大学計算機センターには大変感謝致します。