

ELABORATO 2 - PRODOTTO MATRICE VETTORE PER RIGHE

LO BRUTTO FABIO / MAIONE PAOLO

DEFINIZIONE DEL PROBLEMA

Si vuole progettare un algoritmo in MPI per risolvere il prodotto vettoriale tra una matrice, di dimensioni $M \times N$, e un vettore di reali, di dimensione N , su p processori.

In particolare si utilizza l'infrastruttura S.C.o.P.E. per permettere l'esecuzione del software in un ambiente parallelo.

DESCRIZIONE DELL'ALGORITMO

In particolare le fasi dell'algoritmo, implementato nel file *eLaborato_2.c*, sono:

- 1) Distribuzione per righe della matrice in p processori e distribuzione del vettore: ognuno dei p processori eseguirà il prodotto vettoriale tra la porzione di matrice ricevuta dal processo *root* (cioè quello con rank 0), e il vettore;
- 2) Elaborazione del prodotto vettoriale in parallelo ;
- 3) Aggregazione del vettore calcolata nel processo *root* che determinerà il risultato finale.

A tal proposito sono state utilizzate le primitive fornite da MPI (rispettivamente per la prima fase `MPI_Scatterv()` e per la terza `MPI_Gatherv()`).

Inoltre l'algoritmo progettato comprende anche il caso in cui la dimensione M (cioè il numero di righe della matrice) non sia multipla del numero di processori p a disposizione.

Si è scelto di misurare i tempi di esecuzione nel processo di rank 0 usando la primitiva `MPI_Wtime()` tra la fase 2 e la fase 3 scegliendo il minimo tra 3 misurazioni ripetute.

Infine, si osservi che i controlli di robustezza del software sono stati interamente delegati al processo *root*.

INPUT, OUTPUT E CONDIZIONI DI ERRORE

- **Input:** la matrice e il vettore di cui effettuare il prodotto vettoriale, le loro dimensioni M , N e `dim_vett`.
- **Output:** il vettore risultato dell prodotto vettoriale tra la matrice e il vettore.


```
%un esempio per ciascuna condizione di errore
errori
```

Errore: numero di processori non positivo

```
M63000769@ui-studenti-elaborati/elaborato_2/prodotto_matrice_vettore_righe
```

```
[M63000769@ui-studenti elaborato_2/prodotto_matrice_vettore_righe]$ cat elaborato_2.out
```

```
-----  
This job is allocated on 4 cpu(s)  
Job is running on node(s):  
wn273.scope.unina.it  
wn273.scope.unina.it  
wn273.scope.unina.it  
wn273.scope.unina.it  
-----  
  
PBS: gsub is running on ui-studenti.scope.unina.it  
PBS: originating queue is studenti  
PBS: executing queue is studenti  
PBS: working directory is /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe  
PBS: execution mode is PBS BATCH  
PBS: job identifier is 3921608.torque02.scope.unina.it  
PBS: job name is elaborato_2  
PBS: node file is /var/spool/pbs/aux/3921608.torque02.scope.unina.it  
PBS: current home directory is /homes/DIS/CALCPAR/2019/M63000769  
PBS: PATH = /usr/lib64/openmpi/1.2.7-gcc/bin:/usr/kerberos/bin:/opt/exp_soft/unina.it/intel/composer xe 2013 sp1.3.174/bin/intel64;/opt/exp_soft/  
unina.it/intel/composer xe 2013 sp1.3.174/mpit/bin/intel64;/opt/exp_soft/unina.it/intel/composer xe 2013 sp1.3.174/bin/intel64;/opt/exp_soft/unin  
a.it/intel/composer xe 2013 sp1.3.174/bin/intel64 mic;/opt/exp_soft/unina.it/intel/composer xe 2013 sp1.3.174/debugger/gui/intel64;/opt/d-cache/  
srn/bin;/opt/d-cache/dcqp/bin;/opt/edg/bin;/opt/glite/bin;/opt/globus/bin;/opt/lcg/bin;/usr/local/bin:/bin:/usr/bin:/opt/exp_soft/HADOOP/hadoop-1  
.v.0.3/bin;/opt/exp_soft/unina.it/intel/composerverx/bin/intel64;/opt/exp_soft/unina.it/MFJExpress/mpj-v0_38/bin:/homes/DIS/CALCPAR/2019/M63000769/b  
in  
-----  
Eseguito /usr/lib64/openmpi/1.4-gcc/bin/mpicc -o /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe/elabora  
t_2 /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe/elaborato_2.C  
Eseguito /usr/lib64/openmpi/1.4-gcc/bin/mpixec -machinefile /var/spool/pbs/aux/3921608.torque02.scope.unina.it -np 4 /homes/DIS/CALCPAR/2019/M63  
000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe/elaborato_2  
Errore! Il numero dei processi deve essere positivo  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$  
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
```

```
200      400      600      800     1000     1200     1400     1600     1800
```

100
200
300
400
500
600
700
800
900
1000

1000

100
200
300
400
500
600
700
800
900
1000

1000

Errore: la dimensione di riga della matrice è inferiore al numero di processori

```
M63000769@ui-studenti-elaborati/elaborato_2/prodotto_matrice_vettore_righe
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$ cat elaborato_2.out
-----
This job is allocated on 8 cpu(s)
Job is running on node(s):
wm273.scope.unina.it
wm273.scope.unina.it
wm273.scope.unina.it
wm273.scope.unina.it
wm273.scope.unina.it
wm273.scope.unina.it
-----
PBS: qsub is running on ui-studenti.scope.unina.it
PBS: originating queue is studenti
PBS: executing queue is studenti
PBS: working directory is /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe
PBS: execution mode is PBS BATCH
PBS: job identifier is 3921640.torque02.scope.unina.it
PBS: job name is elaborato_2
PBS: node file is /var/spool/pbs/aux/3921640.torque02.scope.unina.it
PBS: current home directory is /homes/DIS/CALCPAR/2019/M63000769
PBS: PATH = /usr/lib64/openmpi/1.2.7-gcc/bin:/usr/kerberos/bin:/opt/exp_soft/unina.it/intel/composer_xe_2013_spl.3.174/bin/intel64:/opt/exp_soft/unina.it/intel/composer_xe_2013_spl.3.174/mpiir/bin/intel64:/opt/exp_soft/unina.it/intel/composer_xe_2013_spl.3.174/bin/intel64:/opt/exp_soft/unina.it/intel/composer_xe_2013_spl.3.174/debugger/gui/intel64:/opt/d-cache/arm/bin:/opt/d-cache/dcap/bin:/opt/edg/bin:/opt/glite/bin:/opt/globus/bin:/opt/lcg/bin:/usr/local/bin:/bin:/usr/bin:/opt/exp_soft/HADOOP/hadoop-1.0.3/bin:/opt/exp_soft/unina.it/intel/composerxe/bin/intel64:/opt/exp_soft/unina.it/MPExpress/mpj-v0_38/bin:/homes/DIS/CALCPAR/2019/M63000769/bin
-----
Eseguo: /usr/lib64/openmpi/1.4-gcc/bin/mpicc -o /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe/elaborato_2 /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe/elaborato_2.c
Eseguo: /usr/lib64/openmpi/1.4-gcc/bin/mpixec -machinefile /var/spool/pbs/aux/3921640.torque02.scope.unina.it -np 8 /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe/elaborato_2
Errore! La dimensione di riga della matrice è inferiore al numero di processori.
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
```

Errore: le dimensioni non sono coerenti

```
M63000769@ui-studenti-elaborati/elaborato_2/prodotto_matrice_vettore_righe
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$ cat elaborato_2.out
-----
This job is allocated on 4 cpu(s)
Job is running on node(s):
wm273.scope.unina.it
wm273.scope.unina.it
wm273.scope.unina.it
wm273.scope.unina.it
-----
PBS: qsub is running on ui-studenti.scope.unina.it
PBS: originating queue is studenti
PBS: executing queue is studenti
PBS: working directory is /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe
PBS: execution mode is PBS BATCH
PBS: job identifier is 3921659.torque02.scope.unina.it
PBS: job name is elaborato_2
PBS: node file is /var/spool/pbs/aux/3921659.torque02.scope.unina.it
PBS: current home directory is /homes/DIS/CALCPAR/2019/M63000769
PBS: PATH = /usr/lib64/openmpi/1.2.7-gcc/bin:/usr/kerberos/bin:/opt/exp_soft/unina.it/intel/composer_xe_2013_spl.3.174/bin/intel64:/opt/exp_soft/unina.it/intel/composer_xe_2013_spl.3.174/mpiir/bin/intel64:/opt/exp_soft/unina.it/intel/composer_xe_2013_spl.3.174/bin/intel64:/opt/exp_soft/unina.it/intel/composer_xe_2013_spl.3.174/debugger/gui/intel64:/opt/d-cache/arm/bin:/opt/d-cache/dcap/bin:/opt/edg/bin:/opt/glite/bin:/opt/globus/bin:/opt/lcg/bin:/usr/local/bin:/bin:/usr/bin:/opt/exp_soft/HADOOP/hadoop-1.0.3/bin:/opt/exp_soft/unina.it/intel/composerxe/bin/intel64:/opt/exp_soft/unina.it/MPExpress/mpj-v0_38/bin:/homes/DIS/CALCPAR/2019/M63000769/bin
-----
Eseguo: /usr/lib64/openmpi/1.4-gcc/bin/mpicc -o /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe/elaborato_2 /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe/elaborato_2.c
Eseguo: /usr/lib64/openmpi/1.4-gcc/bin/mpixec -machinefile /var/spool/pbs/aux/3921659.torque02.scope.unina.it -np 4 /homes/DIS/CALCPAR/2019/M63000769/elaborati/elaborato_2/prodotto_matrice_vettore_righe/elaborato_2
Errore! Le dimensioni non sono coerenti
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
[M63000769@ui-studenti prodotto_matrice_vettore_righe]$
```

ANALISI DELLE PRESTAZIONI (T(p), S(p), E(p))

Tempo di esecuzione - T(p)

Si è scelto di misurare i tempi di esecuzione nel processo *root* usando la primitiva `MPI_Wtime()`. In particolare l'intervallo di tempo misurato è quello che comprende le fasi 2 e 3 dell'algoritmo prima citate.

Si è scelto inoltre di considerare tre casi diversi che rappresentano le tre possibilità per le dimensioni della matrice: il caso in cui il numero di righe è maggiore numero di colonne, il caso in cui il numero di righe è minore del numero di colonne ed, infine, il caso in cui la matrice è quadrata.

Per ciascuna misurazione corrispondente ai tre casi

- se il numero di righe M è maggiore del numero di colonne N si fa variare M da 5000 a 50000 con N fissato a 1000
- se il numero di righe M è minore del numero di colonne N si fa variare N da 5000 a 50000 con M fissato a 1000
- se la matrice è quadrata si fa variare $M=N$ da 100 a 10000

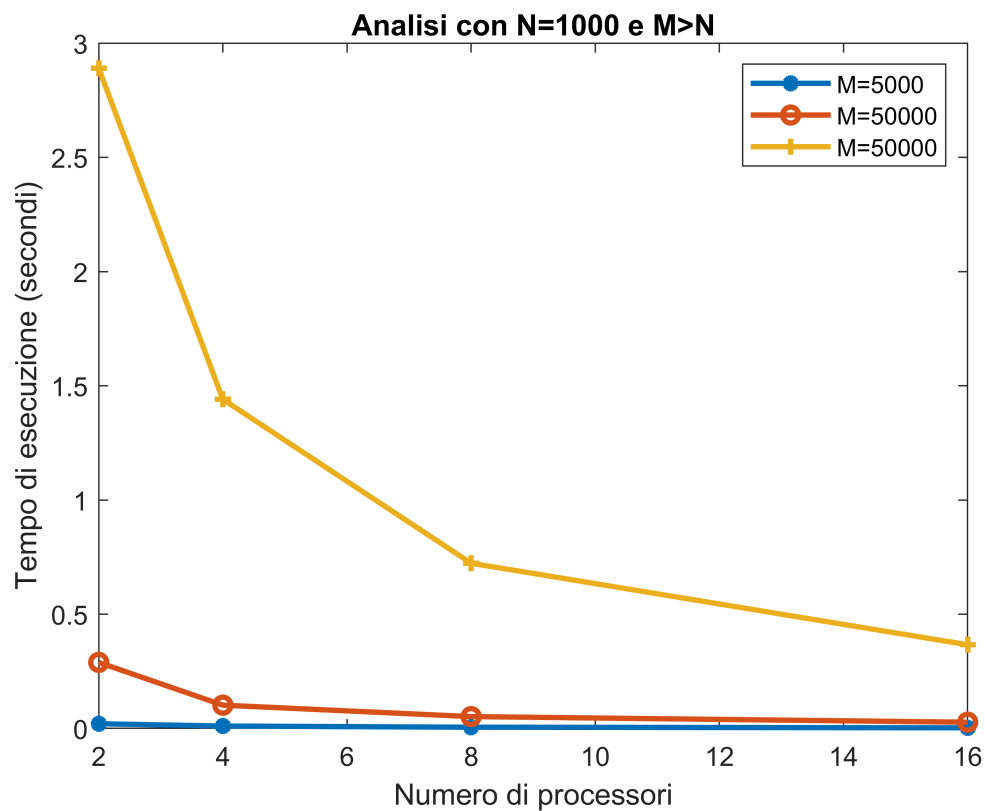
è stato considerato il minimo tra 3 esecuzioni ripetute, eseguite in momenti diversi.

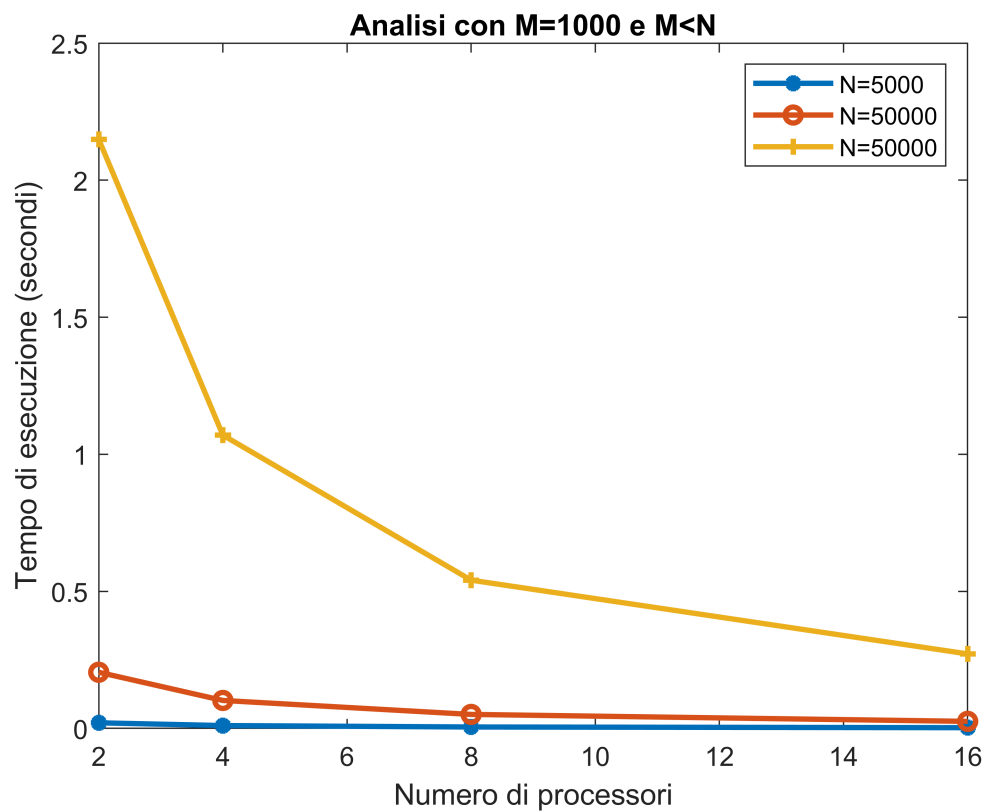
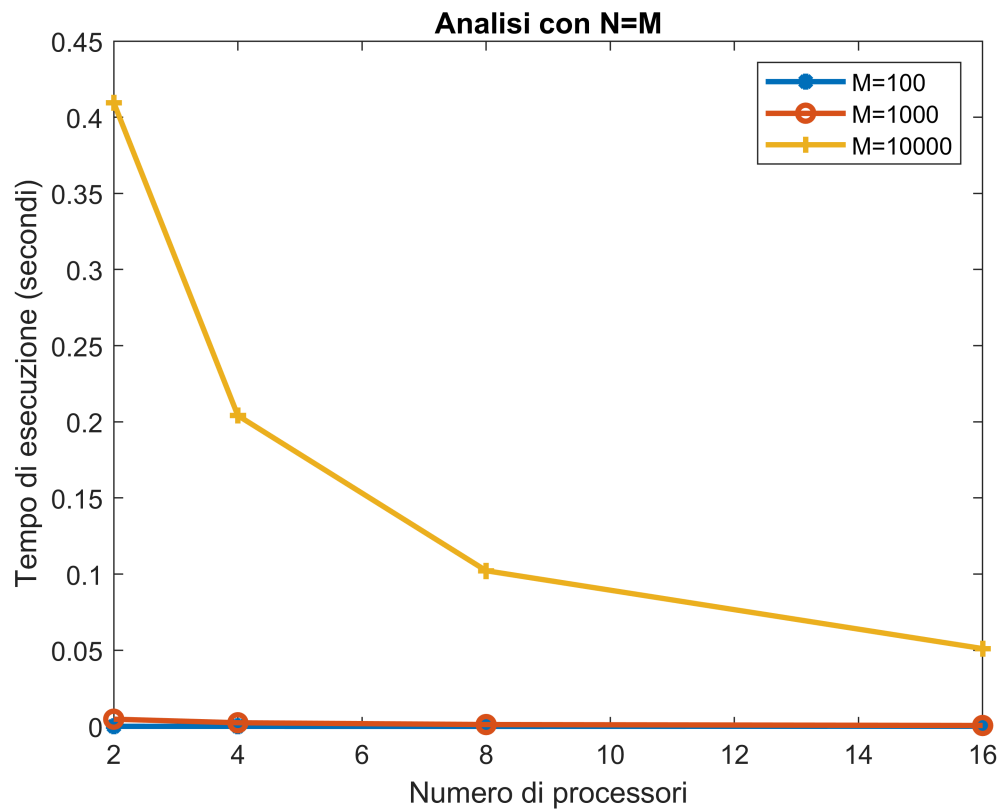
Si noti che nel caso di matrici quadrate si è scelta una diversa configurazione rispetto al caso di matrici sbilanciate data l'impossibilità di eseguire sul cluster il programma con dimensioni 50000 x 50000. Pertanto è parso ragionevole, solo nel caso di matrici quadrate, ridurre la dimensione massima.

I grafici e le tabelle riassumono i risultati ottenuti per tutti i possibili valori di M,N e p.

```
%esecuzione script per tabelle e grafici  
tempi
```

Tempi			
PROVA R>C	C=1000		
	5000	50000	500000
2	0,0203859806060791	0,2891139984130850	2,889853954315180
4	0,0102760791778564	0,1017658710479730	1,440534114837640
8	0,0051968097686768	0,0517168045043945	0,723032951354980
16	0,0026328563690186	0,0271680355072021	0,366197824478149
PROVA R=C			
	100	1000	10000
2	0,0000529289245605	0,0047409534454346	0,409604787826538
4	0,0000331401824951	0,0024030208587646	0,204131841659545
8	0,0000200271606445	0,0012221336364746	0,102236032485961
16	0,0001120567321777	0,0006000316619873	0,051082134246826
PROVA R<C	R=1000		
	5000	50000	500000
2	0,0204110145568848	0,204815864562988	2,149081993103020
4	0,0103540420532227	0,101985931396484	1,069822072982780
8	0,0051629543304443	0,051028013229370	0,540822982788085
16	0,0026011466979980	0,025714159011841	0,271781930923461





Per considerazioni più di dettaglio su questi risultati si rimanda alla sezione Conclusioni.

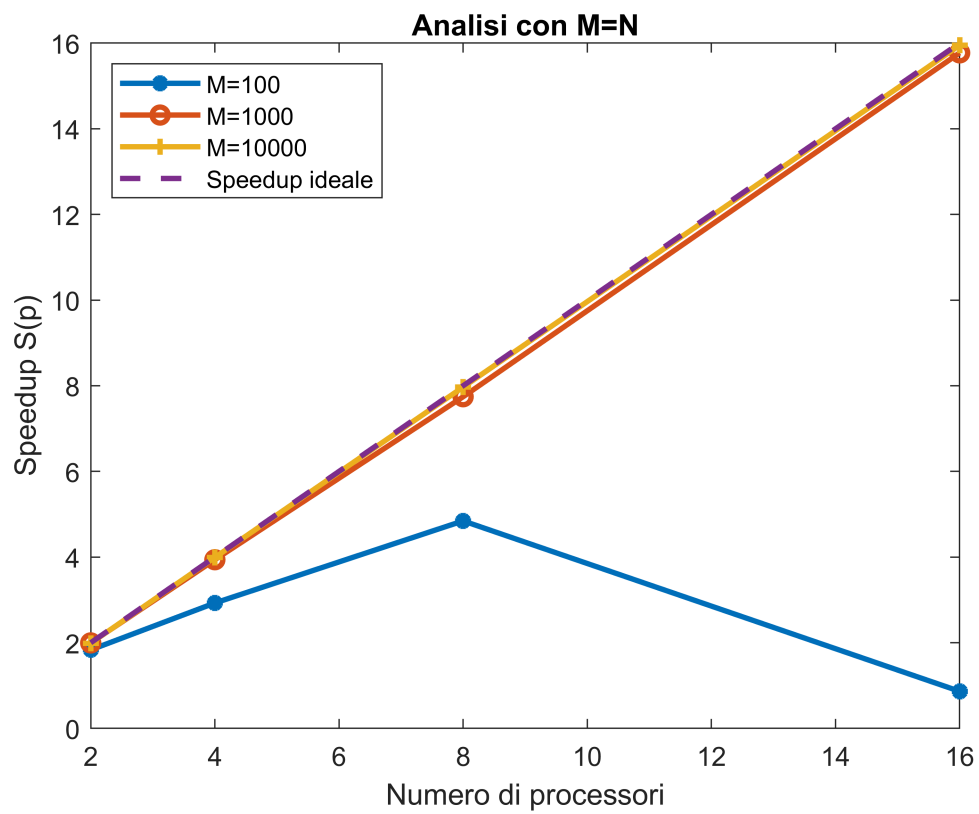
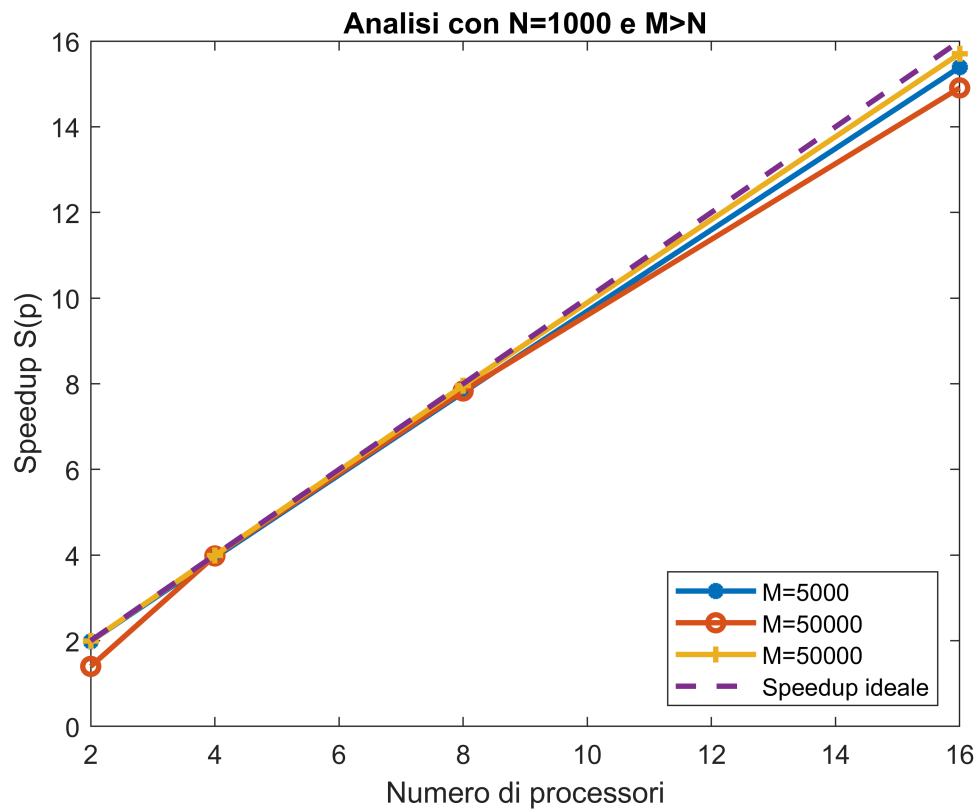
Speed up ed Efficienza - S(p) ed E(p)

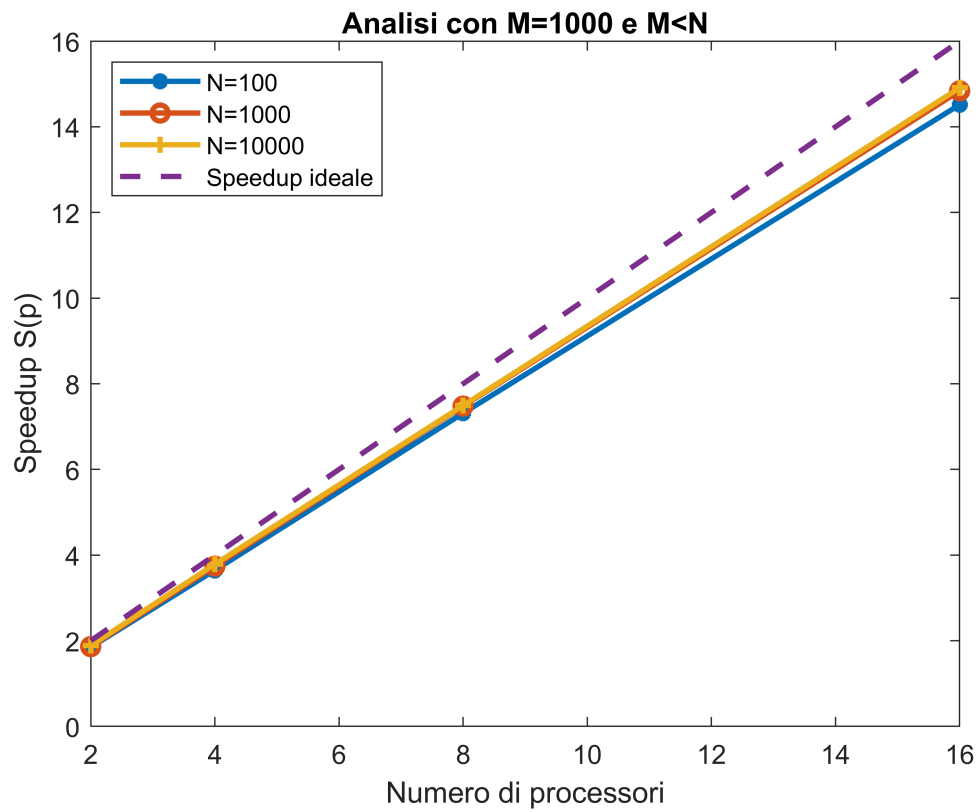
Si è calcolato, inoltre, il tempo di riferimento T(1) che corrisponde al tempo di esecuzione su un unico processore.

A partire dai tempi misurati nella sezione precedente e da T(1) è stato calcolato lo speed-up al variare di M, N e p.

%esecuzione script per tabelle e grafici
speedup

Speed up			
PROVA R>C	C=1000		
	5000	50000	500000
2	1,9877	1,4009	1,9904
4	3,9432	3,9799	3,9929
8	7,7973	7,8315	7,9553
16	15,3906	14,9079	15,7071
PROVA R=C			
	100	1000	10000
2	1,8333	1,9962	1,9892
4	2,9281	3,9384	3,9914
8	4,8452	7,7439	7,9695
16	0,8660	15,7725	15,9503
PROVA R<C	R=1000		
	5000	50000	500000
2	1,8499	1,8629	1,8855
4	3,6467	3,7413	3,7876
8	7,3133	7,4774	7,4924
16	14,5160	14,8384	14,9093

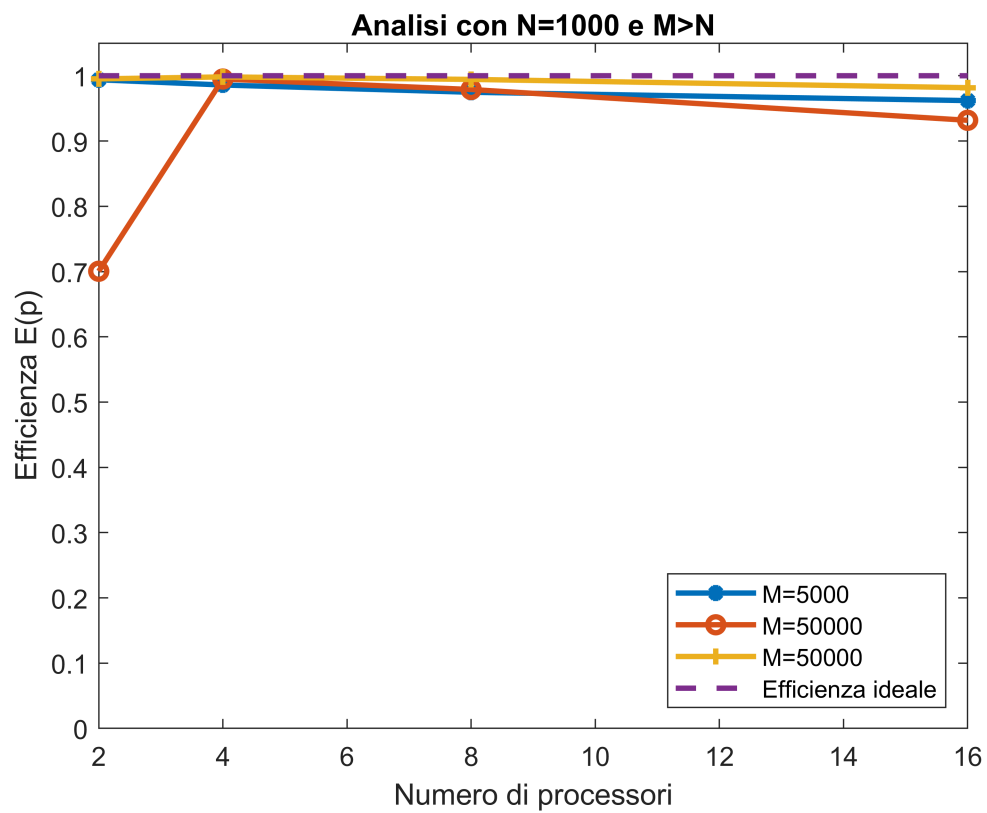


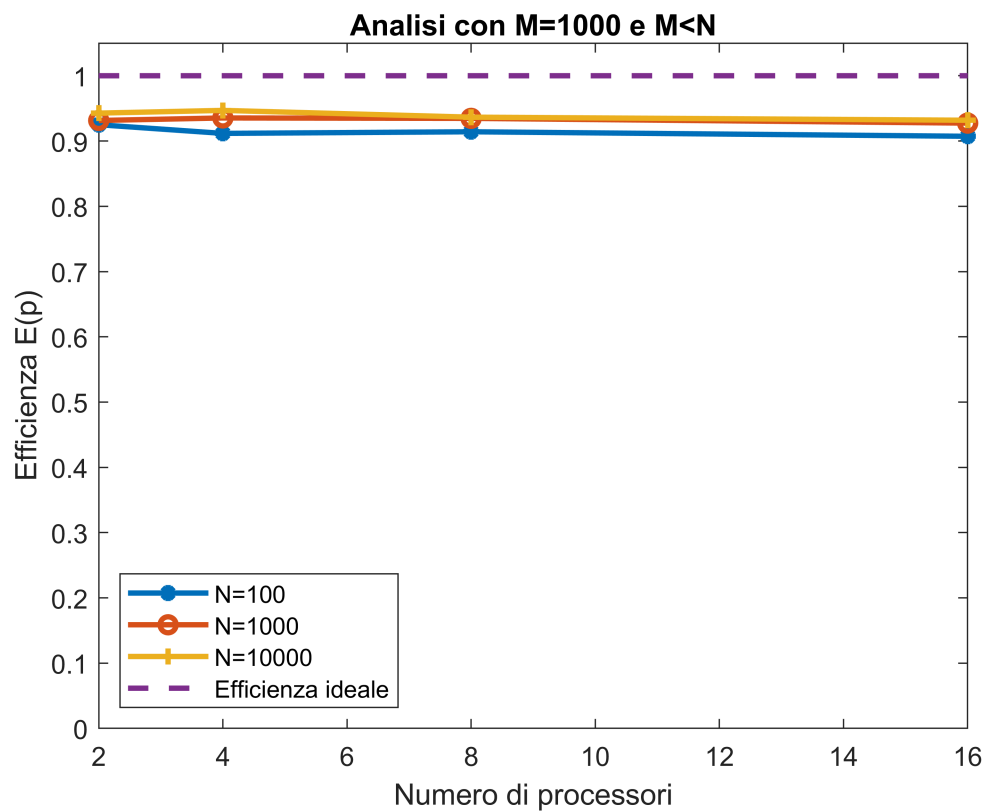
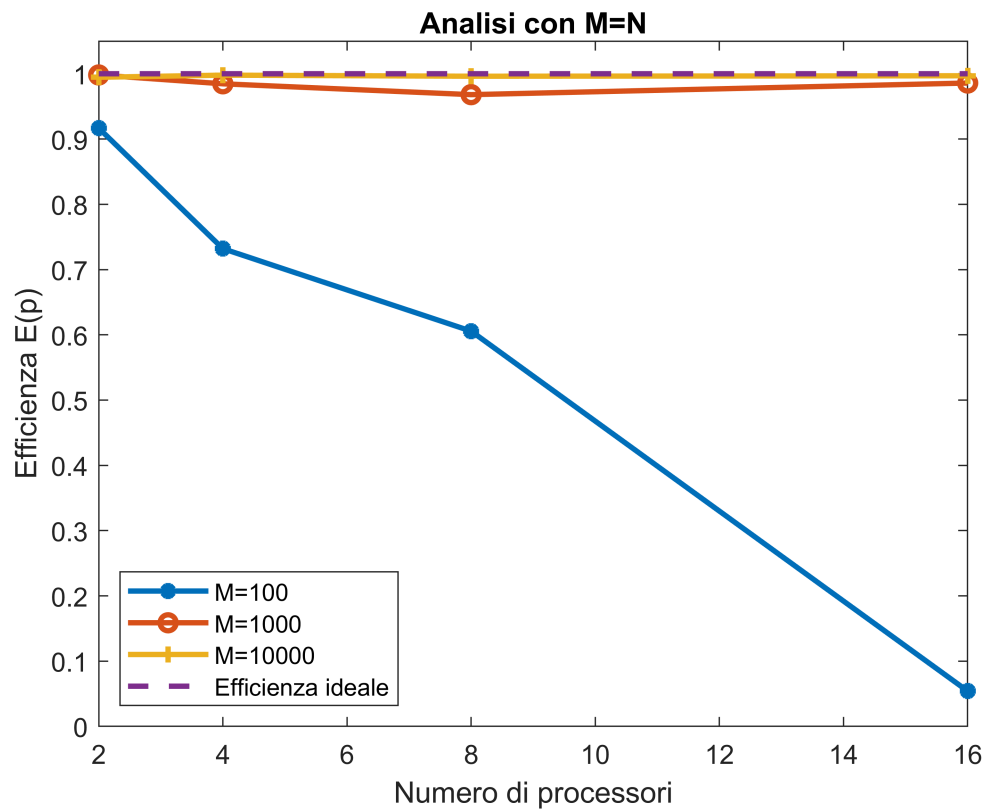


Infine si è calcolata l'efficienza rapportando lo speed-up $S(p)$ al numero di processori p .

```
%esecuzione script per tabelle e grafici  
efficienza
```

Efficienza			
PROVA R>C	C=1000		
	5000	50000	500000
2	0,9938	0,7004	0,9952
4	0,9858	0,9950	0,9982
8	0,9747	0,9789	0,9944
16	0,9619	0,9317	0,9817
PROVA R=C			
	100	1000	10000
2	0,9167	0,9981	0,9946
4	0,7320	0,9846	0,9979
8	0,6057	0,9680	0,9962
16	0,0541	0,9858	0,9969
PROVA R<C	R=1000		
	5000	50000	500000
2	0,9249	0,9315	0,9427
4	0,9117	0,9353	0,9469
8	0,9142	0,9347	0,9366
16	0,9072	0,9274	0,9318





Conclusioni

Dai grafici e dalle tabelle appena presentate si possono trarre considerazioni diverse nei tre casi prima citati:

- se $M > N$ o $M < N$ l'efficienza migliore si ha per 2 processori quando $M=5000$. Se $M=50000$ o 500000 invece il numero ottimo di processori risulta pari a 4. Tuttavia l'efficienza risulta complessivamente maggiore nella configurazione in cui $M > N$.
- se $M=N$ l'efficienza migliore si ha per 2 processori per M fino a 1000 mentre è 4 quando $M=10000$. In particolare si osservi che per $M=100$ l'efficienza degrada rapidamente all'aumentare del numero di processori date le dimensioni ridotte.

Si possono fare ulteriori considerazioni notando che, per una dimensione fissata, l'efficienza peggiora dopo un certo valore di p (ciò verifica sperimentalmente la legge di Amdahl), e che, in generale, all'aumentare sia della dimensione del problema che di p , l'efficienza migliora o non peggiora come nel caso di $M=N$ (verificando la legge di Gustafson).

Analoghe considerazioni per i tempi e lo speedup.

ANALISI DELL' ACCURATEZZA

Confrontando il risultato ottenuto sul cluster Scope e quello ottenuto su MATLAB si ottiene il seguente errore relativo (tramite il comando `norm`), fissando a 4 il numero di processori con dimensione della matrice pari a 7×8 .

```
%esecuzione script per i test di accuratezza
accuratezza
```

```
errore_relativo = 4.3402e-06
```