Blocked In-Place Rectangular Transpose

Fred Gustavson
Umea Univ. || class
May 10, 2007

Main Idea

- Combine Cache Blocking with Point In-Place Transpose on a very tiny matrix
  - use of SB Format is the key idea
  - CM -> SB -> SB^T -> CM
- Block In-Place Transpose is Very Fast relative to Point In-Place Transpose
- CM <-> SB uses fast vector In-Place Alg.

Summary or Overview

- A is M by N.
- M=m*NB & N=n*NB
- CM -> SB by vector IP transpose
- SB <-> SB^T by block IP transpose
  - use point IP transpose on m by n A1= SB A
- SB^T -> CM by vector IP transpose

Vector In-Place Xpose or CM<->SB

- Let A be M by NB with M=m*NB
- view A as m by NB A1 with each a1(i,j) being a column vector of size NB
- apply point IP transpose to A1 to get A2
- A2 is m order NB SB's concatenated
- Apply above subroutine n=N/NB times

Where does the Speed Come From

- Data moved in blocks and vectors gives a 10 to 100 times performance gain
  - uses stride one processing; every line gets fully used when it enters L1 and streaming by algorithmic / automatic pre-fetching works
- SMP parallelism is easy to implement
  - disjoint cycle structure
  - long cycles can be broken into pieces

Other Matrix Layouts

- can block transform (in-place) any permutation that can be described by a compact functional description
  - includes all common matrix data layouts
    - standard CM / RM rectangular arrays
    - standard CM / RM triangular arrays
    - standard packed format

A is 500 by 700 in CM order

- CM A has LDA = 500
- A has 7 column swaths: 500 by 100 each
- A1 is 5 by 100 matrix of vectors
- In-place transpose with q = 499
- repeat above 6 more times
- A is now in SB format of size 5 by 7

Details of CM to SB Vector

- 0 and m*n -1 = 499 are singleton cycles
- 499 is prime and # d = 2; 1 & 499
- q = m*n – 1 is the mod value
- for problem 499, phi = 498 & cl = 249; leaders are 1, 2
- for problem 1, phi = 1 & cl = 1 at 499

Details of SB to SB^T

- q = 5*7 – 1 = 34 = 2*17
- q = sum over divisors of phi
  - #d = 4; 34, 17, 2, 1; phi's = 16, 16, 1, 1
- #d problems gives cycles of length 16, 16, 1, 1 starting at 1, 2, 17, 34

Details of SB^T to CM

- m = 100, n = 7, q = 699 = 3*233
- # d = 4; 699, 233, 3, 1; phi's 464, 232, 2, 1
- cl's are 166, 166, 1, 1
- leaders are 1, 2, 5, 10; 3, 9; 233, 466; 699

The 500 by 700 A as a point matrix

- q = m*n – 1 = 349,999 = 13**2*19*109
- # d's = 3*2*2 = 12:
- sum of phi(d) = q
- twelve phi's are 303264, 23328, 16848, 2808, 1944, 1296, 216, 156, 108, 18, 12, 1
- twelve cl's are 468, 36, 156, 468, 18, 12, 36, 156, 6, 9, 12, 1
- ratio's give # of leaders: 648, 648, 108, 6,108, 108, 6, 1, 18, 2, 1, 1: sum = 1655

500 by 700 A as point matrix

- hand-out has cycle of length 12 at ij=247
  - ij is (247,0) element of A; next element in cycle is mod(247*700,q); 247 | q so we get cycle is i <- mod(700*i,1417) :
  - 247*(1, 700, 1135, 980, 172, 1372, 1091, 1244, 762, 608, 500, 1) mod (q)