Suggestion Answer for Homework 1, CS421

September 22, 1999

1. Six ways to compute $C = A*B$ are obtained by different ordering of the triply nested loop.

row-by-row

```
for i = 1:n
   for j = 1:n
      for k = 1:n
         C(i,j) = C(i,j) + A(i,k) * B(k,j)
      end
   end
end
```

column-by-column

```
for j = 1:n
   for i = 1:n
      for k = 1:n
         C(i,j) = C(i,j) + A(i,k) * B(k,j)
      end
   end
end
```

The (i-k-j) ordering: update $C(i,:)$ by $A(i,k) * B(k,:)$
The (j-k-i) ordering: update $C(:,j)$ by $B(k,j) * A(:,k)$
The (k-i-j) ordering: update $C(i,j)$ by inner product of $A(i,:)$ and $B(:,j)$
The (k-j-i) ordering: similar to (k-i-j) ordering.

2.$W = AV = XY^TV = X(Y^TV)$ where X,Y,V are column vector.
First compute $\alpha = Y^TV$ which needs $2n$ flops.
Then compute $W = \alpha * X$ which needs n flops.

3.Denote T_1,T_2,...T_n to be the column vectors of T. Let us consider linear equations $TX = 0$. Because $T(i,i) \neq 0$, $X = 0$ is the unique solution of this linear system. So $\sum_{i=1}^{n} x_i T_i = 0$ only has trival solution. By the definition of linear independency, $\{T_i | i = 1, 2, ...n\}$ are linearly independent. Thus $rank(T) = n$.

If there is some $T(j, j) \neq 0$, let j_1 be the first zero entry on the diagonal. Then we can find constant $\alpha_1, ...\alpha_{j_1-1}$ such that $T(j_1, :) = \Sigma_{i=1}^{j_1-1} \alpha_i T(i, :)$. So $rank(T) < n$.

4. This algorithm computes Gauss transfomrs M_1, ... M_{n-1} and interchange permutations E_1, ...E_{n-1} such that $M_{n-1}E_{n-1}...M_1E_1A = U$ is upper triangular. No multiplier is bigger than 1 in absolute value. $A(1 : k, k)$ is overwritten by $U(1 : k, k), k = 1 : n$. $A(k + 1 : n, k)$ is overwritten by $-M_k(k + 1 : n, k), k = 1 : n - 1$.

for k= 1:n-1
  find $\mu = max|A(i : n, i)|$
  switch $A(k, 1 : n)$ with $A(\mu, 1 : n)$
  $E_k = I_{k,\mu}$ where $I_{k\mu}$ is an interchange permutation involving row k and a row μ.
  if $A(k, k) \neq 0$
    rows = k+1:n
    $A(rows, k) = A(rows, k)/A(k, k)$
    $A(rows, rows) = A(rows, rows) - A(rows, k)A(k, rows)$
  end
end

We now have $M_{n-1}E_{n-1} \cdots M_1E_1A = U$.
Let $\tilde{M}_{n-1} = M_{n-1}$ and $\tilde{M}_k = E_{n-1} \cdots E_{k+1}M_kE_{k+1} \cdots E_{n-1}$ $k \leq n - 2$.
Then $\tilde{M}_{n-1} \cdots \tilde{M}_1PA = U$ where $P = E_{n-1} \cdots E_1$.
Since $M_k = I - \tau^k e_k^T$, $L(k+1 : n, k) = g(k+1 : n)$ where $g = E_{n-1} \cdots E_{k+1}\tau^k$. Thus each $\tilde{M}_k$ is a Gauss transform with Gauss vector $\tilde{\tau}^k = E_{n-1} \cdots E_{k+1}\tau^k$. So upon completion, $A(i, j)$ houses $L(i, j)$ for all $i > j$.

5. Denote S to be the set of tuples $(i, j, Value_{i,j})$.

$result$ = zeros(n,1)
while there are more 3-tuples in S
  $result(i) = result(i) + Value_{i,j} * x(j)$
  remove current $(i, j, Value_{i,j})$ from S
endwhile

flops = $O(\sum_{i=1}^{n} r_i) = O(\sum_{j=1}^{n} c_j)$

6.By board cholesky, we can continue cholesky factorization.
flops= $O((m + n)^3 - n^3) = O(mn^2)$. Because $m \ll n$, flops = $O(n^2)$.

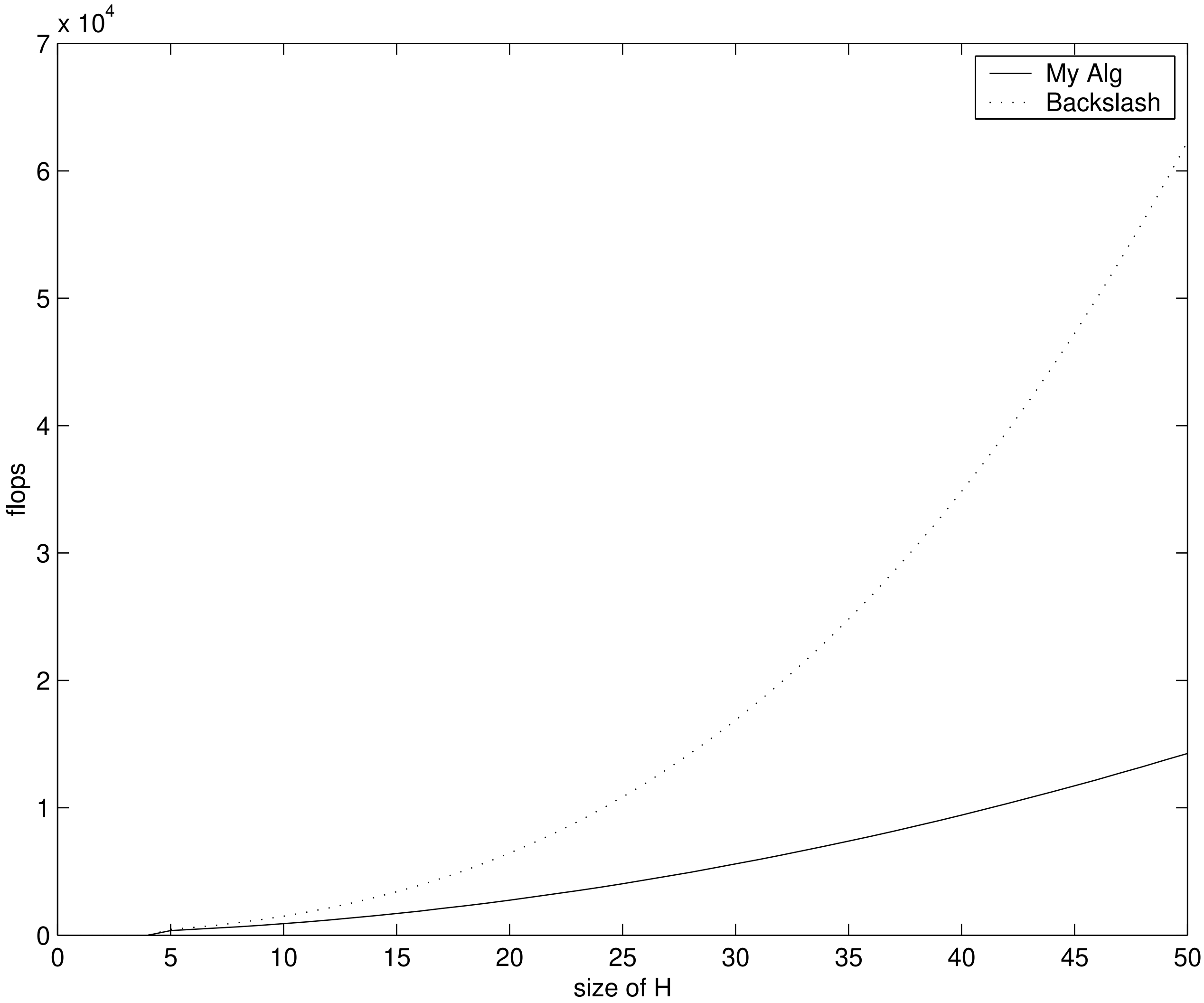

Figure 1: number of flops

Code:

```
...
for k = n+1:n+m
    L(k,1:k-1) = (L(1:k-1,1:k-1)  M(1:k-1,k))';
    L(k,k) = sqrt(M(k,k) - L(k,1:k-1)* L(k,1:k-1)');
end
...
```

7. Let

$$M = \begin{bmatrix} -W & A \\ A^T & H \end{bmatrix} \quad M_1 = \begin{bmatrix} I & 0 \\ A^T W^{-1} & I \end{bmatrix}$$

Then

$$M_1 M = \begin{bmatrix} -W & A \\ 0 & H + A^T W^{-1} A \end{bmatrix}$$

Instead of solving $(H + A^T W^{-1} A)z = r$, we consider the problem

$$\begin{bmatrix} -W & A \\ 0 & H + A^T W^{-1} A \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ r \end{bmatrix}$$

Because

$$M_1^{-1} = \begin{bmatrix} I & 0 \\ -A^T W^{-1} & I \end{bmatrix}$$

we only need to solve the problem

$$M \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} -W & A \\ A^T & H \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = M_1^{-1} \begin{bmatrix} 0 \\ r \end{bmatrix} = \begin{bmatrix} 0 \\ r \end{bmatrix}$$

If the property of M is very good, such as, being a sparse matrix, we can find the solution of $M[y, z]^T = [0, r]^T$ efficiently. Then we can get the solution z directly. Following this algorithm we need not find $H + A^T W^{-1} A$, $A^T W^{-1} A$ and W^{-1}.