AMSC/CMSC 661 Scientific Computing II
Spring 2010
The Fast Multipole Method
Dianne P. O'Leary

The Fast Multipole Method

References:

Idea due to Greengard and Rokhlin.

Xiaobai Sun and Nikos P. Pitsianis, A matrix version of the fast multipole method, SIAM Review 43 (2001) 289-300.

Nail A Gumerov and Ramani Duraiswami, Fast Multipole Methods for the Helmholtz Equation in Three Dimensions (The Elsevier Electromagnetism Series)

What is the FMM?

In many applications, it is important to compute, for example,

- the gravitational potential arising from a distribution of masses
- electrostatic potential arising from a distribution of charges

Picture. We'll talk about charges, for definiteness.

Suppose that the charge on source particle j, which is located at position s_j, is q_j, $j = 1, \ldots, n$. Then to compute the potential p_k at target particle k, located at position t_k, we compute

$$p_k = \sum_{j=1}^{n} \frac{q_j}{\|t_k - s_j\|^\beta}.$$

$k = 1, \ldots, m$.

For notational convenience, we take $\beta = 1$, but it really doesn't matter.

Notice that computing all of the potentials is just a matrix-vector product

$$\mathbf{p} = \mathbf{A}\mathbf{q}$$

where

$$a_{kj} = \frac{1}{\|t_k - s_j\|}.$$

The Fast Multipole Method (FMM) provides a fast way to approximately evaluate $\mathbf{Aq}$.

What is the underlying idea?

$$\mathbf{p} = \mathbf{A}\mathbf{q}$$

where

$$a_{kj} = \frac{1}{\|t_k - s_j\|}.$$

- Suppose that many of the source particles were located at the same place. Then several columns of the matrix $\mathbf{A}$ would be identical, and we could compress the matrix to one with fewer columns by adding the corresponding elements of $\mathbf{q}$.
- Similarly, if several of the target particles were at one location, then we could delete the redundant rows of the matrix $\mathbf{A}$.

In either case, we end up with an equivalent problem with a smaller matrix and therefore a faster matrix-vector product.

An approximation

The FMM is built on the idea of approximating the matrix-vector product by moving source particles that are close to each other, and far from the target, to their centroid, and doing the same with the targets.

This means that the matrix $\mathbf{A}$ is replaced by a matrix $\mathbf{A}_r$ of rank $r < m, n$, and the cost of matrix-vector product is reduced from $O(mn)$ to $O(mr + nr)$.

If the approximation is not accurate enough, a correction term can be computed.

All of this can be done recursively.

The approximation for single clusters of sources and targets

Theorem: Suppose we have a set of sources, centered at s_c, and a set of targets, centered at t_c, with $\alpha < 1$ chosen so that

$$\max_j \|s_j - s_c\| + \max_k \|t_k - t_c\| \leq \alpha \|t_c - s_c\|.$$

Then given any integer $p \geq 0$

$$\mathbf{A} = \mathbf{A}_r + \mathbf{A} \odot \mathbf{E}$$

where

- $\mathbf{A}_r$ has rank at most $r = (p+1)(2p+1)$.
- $\odot$ denotes the Hadamard matrix product

$$(\mathbf{A} \odot \mathbf{B})_{kj} = a_{kj} b_{kj}.$$

- The elements of $\mathbf{E}$ are bounded by

$$|e_{kj}| \leq \frac{1+\alpha}{1-\alpha}\alpha^{p-1}.$$

Note: We can make this arbitrarily accurate by choosing r large enough. ($r = \min(m, n)$ gives the exact result)

What if the sources and targets are interspersed?

Then partition them!

In 1-d, for example, divide the sources into those centered around $s_o < 0$ and those not, and divide the targets in a similar way. Then

$$\mathbf{A} = \mathbf{A}_f + \mathbf{A}_n$$

where

$$\mathbf{A}_f = \begin{bmatrix} 0 & \mathbf{A}_{o,n} \\ \mathbf{A}_{n,o} & 0 \end{bmatrix}$$

represents the far-field interactions and

$$\mathbf{A}_n = \begin{bmatrix} \mathbf{A}_{o,o} & 0 \\ 0 & \mathbf{A}_{n,n} \end{bmatrix}$$

represents the near-field interactions.

Then multiplication by $\mathbf{A}_f$ can be done by FMM.

The recursion

We are left with the problem of forming $\mathbf{A}_{o,o}\mathbf{q}_o$ and $\mathbf{A}_{n,n}\mathbf{q}_n$. These are two smaller problems of the same form, so we just recurse!

One final trick

If the original sources and targets are located at the mesh points of a grid, then the matrix $\mathbf{A}$ has Toeplitz or block-Toeplitz structure, and the multiplication can be done very quickly using FFTs.

The work

The FMM has number of multiplications and additions proportional to $O(n \log_2 n)$ (when $n \geq m$). This is a great savings over the $O(mn)$ count of the original algorithm.

A connection with PDEs

Green's functions can be interpreted as charge distributions.