# Data Sparse Matrix Computations - Lecture 4

Scribes: John Ryan, Paul Upchurch

September 19, 2017

## Contents



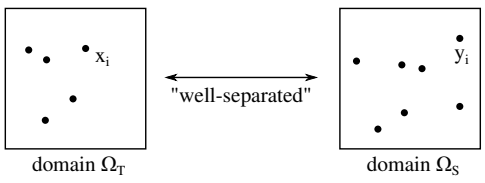
## 1 Fast Multipole Method Continued

### Low-rank assumption

Last time we worked on solutions to N-body problems in the form

$$\phi(x_i) = \sum_{j=1}^{N} K(x_i, x_j) q_j \tag{1}$$

Specifically, let's say we have domains $\Omega_S$ of $N$ source points and $\Omega_T$ of $M$ target points, and these domains are "well-separated" (we will formalize this in section 3). Our goal is to compute the influence of all source points onto target points.




Let the $M \times N$ matrix $[K]_{ij} = K(x_i, y_j)$ and assume it is approximately low-rank, so that $K \approx UV^T$ with $U$ of size $M \times P$ and $V$ of size $N \times P$. If $P$ is small then we can efficiently compute the effect of many points in the source domain on points in the target domain.

This low rank assumption is the same as saying that we can represent $K$ with a function of $x$ only and a function of $y$ only

$$K(x, y) \approx \sum_{l=1}^{P} u_l(x) v_l(y)$$

when $x, y$ are far apart.

## 2 Notation

For the following discussion of the FMM we first need to introduce notation for indexing with respect to subsets of $\Omega$. We will need a clean way to sum over a set of points, $\{x_i\}_{i=1}^{N}$ $x_i \in \mathbb{R}^2$,

which lie in box-like subregions of $\mathbb{R}^2$. If we define $\Omega_T \subset \Omega$ then let $I_T$ represent the indices of all of the points in the domain $\Omega_T$. In particular,

$$i \in I_T \iff x_i \in \Omega_T$$

Then we can rewrite Equation (1) as

$$\Phi(x_i) = \sum_{j \in I_S} K(x_i, y_j) q_j$$

Furthermore, we will use these indices with matrices and vectors. Recall that the linear algebra form of the N-body problem is

$$\Phi = Kq$$

We can use $I_S, I_T$ for domains $\Omega_S, \Omega_T$, respectively, to write the statement

$$\Phi_{\Omega_S}(I_T) = K(I_T, I_S) q(I_S)$$

which means that we assign to a portion of $\Phi_{\Omega_S}$ the matrix-vector product of portions of $K$ and portions of $q$. The portions do not need to be contiguous. The subscript $\Omega_S$ is a reminder that this computation leaves out sources outside $\Omega_S$. In practice, we can avoid forming the entire matrix $K$, which has $O(N^2)$ terms.

# 3 Formalize well-separated

A key assumption of the FMM is that the source and target domains are well-separated. We will formalize what we mean by well-separated. Consequently, we will also show how to factor a kernel into a low-rank approximation which has a known, prescribed accuracy.

Let us take the kernel to be the fundamental solution to Laplace's equation in two dimensions. This is (leaving out the constant $\frac{1}{2\pi}$ factor for brevity)

$$K(\mathbf{x}_i, \mathbf{x}_j) = \begin{cases} 0 & \mathbf{x}_i = \mathbf{x}_j \\ \log\left(|\mathbf{x}_i - \mathbf{x}_j|\right) & \mathbf{x}_i \neq \mathbf{x}_j \end{cases}$$

We can equate a 2D vector $\mathbf{x}$ with a complex number $x$. By $\log(|\mathbf{x}-\mathbf{y}|) = Real(\log(x-y))$ we can drop the mod inside the log. This makes Taylor expanding (which comes up in the approximation) easier, hence let's stick with complex numbers.

Our $x$'s will be in $\Omega_T$ and $y$'s in $\Omega_S$. Say these domains are boxes. The center of $\Omega_T$ will be called $c_T$ and the center of $\Omega_S$ will be called $c_S$.

$$\begin{aligned} K(x, y) &\approx \log(x - y) \\ &= \log\left((x - c_S) - (y - c_S)\right) \\ &= \log(x - c_S) + \log\left(1 - \frac{y - c_S}{x - c_S}\right) \end{aligned}$$

We can approximate this well if $\frac{|y-c_S|}{|x-c_S|}$ is less than 1 (remember the geometric interpretation, which is that $y$ is close to the center of its domain, but $x$ is far away). After a Taylor expansion of the second term, we have

$$= \log(x - c_S) - \sum_{l=1}^{\infty} \frac{1}{l} \frac{(y - c_S)^l}{(x - c_S)^l} \tag{2}$$

### Low-rank approximation and error analysis

By truncating (2) to $P$ terms we have the low-rank approximation

$$\log(x - c_S) - \sum_{l=1}^{P} \frac{1}{l} \frac{(y - c_S)^l}{(x - c_S)^l} \tag{3}$$

whose error is bounded by

$$\left|\sum_{l=P+1}^{\infty} \frac{1}{l}\frac{(y-c_S)^l}{(x-c_S)^l}\right|$$
$$\leq \frac{1}{P+1}\left|\sum_{l=P+1}^{\infty} \frac{(y-c_S)^l}{(x-c_S)^l}\right|$$
$$\leq \frac{1}{P+1}\sum_{l=P+1}^{\infty}\left|\frac{y-c_S}{x-c_S}\right|^l$$

Let $\gamma = |\frac{y-c_s}{x-c_s}|$ and pull out a factor to make a geometric series, which leads to

$$= \frac{1}{P+1}\gamma^{P+1}\frac{1}{1-\gamma}$$

The first two terms depend on $P$ and decay. The last term is a constant. Hence, we can choose $P$ to make this error as small as we want.

Let's just use this approximation when $\gamma < \frac{1}{2}$. We'll discuss this more later.

# 4 Multipole expansion

We have a low-rank way to compute $K$ efficiently but we still need to accumulate the effects of all source points onto each target point. A set of coefficients, $\omega_l^S$, will compactly represent the low-rank contributions of points in source box $\Omega_S$ onto any point which is well-separated.

$$\omega_l^S = \begin{cases} \sum_{j\in I_S} q_j & l = 0 \\ \sum_{j\in I_S} \frac{-1}{l}(y_j - c_S)^l q_j & l \in \{1, 2, \ldots, P\} \end{cases} \tag{4}$$

These coefficients are independent of the target domain so they can be reused many times.

# 5 Local Expansion

Now, substitute the low-rank approximation (3), expressed in terms of coefficients (4), into a summation of (1) over all source points for a target point $x \in \Omega_T$,

$$\phi(x) = \sum_{j\in I_S} K(x, y_j)q_j$$
$$\approx \sum_{j\in I_S} \log(x - c_S)q_j + \sum_{j\in I_S}\sum_{l=1}^{P}\frac{-1}{l}\left(\frac{y_j - c_S}{x - c_S}\right)^l q_j$$
$$= \log(x - c_S)\omega_0^S + \sum_{l=1}^{P}\left(\frac{1}{x - c_S}\right)^l \omega_l^S \tag{5}$$

This function is harmonic on the target domain, which is to say it solves Laplace's equation (this follows from the definition of the kernel as the fundamental solution). Thus it has a nice Taylor expansion.

**Exercise:** Show that the Taylor expansion of (5) around box center $c_T$ is

$$\phi(x) = \sum_{l=0}^{\infty}(x - c_T)^l v_l^T \tag{6}$$

Where $v_l^T$ are some coefficients defined below. You may guess now that there is a fair bit of analytical work involved in the design of the FMM.

$$v_0^T = \omega_0^S \log(c_T - c_S) + \sum_{l=1}^{\infty}\omega_l^S \frac{(-1)^l}{(c_S - c_T)^l} \tag{7}$$

$$v_l^T = -\omega_0^S \frac{1}{l(c_S - c_T)^l} + \sum_{j=1}^{\infty} \omega_j^S (-1)^j \binom{j+l-1}{j-1} \frac{1}{(c_S - c_T)^{l+j}}$$

These sums will get truncated, much like in the earlier Taylor expansion. (Scribe's solution at bottom).

# 6 Back to Linear Algebra

We've constructed a linear operator, so that

$$\omega^S = T_S^M q(I_s)$$

See equation (4). $T_S^M$ is a $P+1$ by $N$ matrix. The $S$ denotes that it depends only on the source domain, and the $M$ denotes that it returns the multipole coefficients.

The $v$ coefficients are a way to take multipole coefficients and give local coordinates. (The terminology in the literature is inconsistent so you may see the coefficients named outgoing/incoming, etc.)

After truncation of (7) we have

$$v_T = T_{T,S}^{M2L} \omega^S$$

where $M2L$ means "multipole to local", and the matrix depends on both domains $T$ and $S$. This matrix is $P+1$ by $P+1$.

Finally we have a local expansion.

$$\phi(I_T) = T_T^L v_T$$

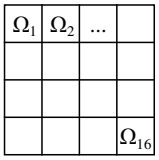
from (6). This matrix is $M$ by $P+1$.

All of this is to say that

$$K(I_T, K_S) = T^L T^{M2L} T_s^M$$

# 7 Uniformly distributed points

Thus far we have not said anything about how to divide $\Omega$ into boxes. Let's consider the case where our points are nearly uniformly distributed inside a 2D domain. Divide $\Omega$ into a grid of $m$ by $m$ boxes $\Omega_1, \Omega_2, ..., \Omega_{m^2}$.



domain Ω

We can write

$$\phi(I_1) = \sum_{l=1}^{m^2} K(I_1, I_l) q(I_l)$$

for points in box $\Omega_1$. Our low-rank approximation framework is not valid for all source boxes. For example, it may be valid for $l = 16$ but not for $l = 2$ (refer to figure) since points in neighboring boxes may not satisfy $\gamma < \frac{1}{2}$.

We will categorize the boxes based on adjacency. For box $\Omega_i$, we define $L_i^n$ as the neighbors of that box (i.e., the boxes that share an edge or corner with $\Omega_i$ in the grid). Also, we define $L_i^f$ to be all far boxes that are not $\Omega_i$ and are not in the neighbor list of $\Omega_i$. Then

$$\Phi(I_i) = \underbrace{K(I_i, I_i) q(I_i)}_{\text{not low-rank}} + \underbrace{\sum_{j \in L_i^n} K(I_i, I_j) q(I_j)}_{\text{not low-rank}} + \underbrace{\sum_{j \in L_i^f} K(I_i, I_j) q(I_j)}_{\text{low-rank}}$$

We want to pick an $m$ so that the cardinality of the neighbor-list of any box is bounded by a constant. Thus, the second sum is likely to be the most computationally expensive.

This suggests a simple algorithm—loop over the boxes, compute $\Phi(I_i)$ in these pieces, use low-rank approximations for the second sum and exact computations for the rest.

- Loop over all boxes and compute multipole expansion coefficients.
- Compute all $T^{M2L}$ operators for almost all pairs of boxes. (It is not needed for neighbors, but that's a small-sized list.)
- Loop over boxes and compute local expansions plus neighbors.

If $m$ is optimal ($m^2 \approx N^{2/3}$), then for large $N$ the middle step costs $\mathcal{O}(N^{4/3})$.

## Next lecture

From here, we will next discuss the case when the points are not uniformly distributed and describe the complete FMM, which is a linear time algorithm.

## Solution to exercise

By definition of Taylor Series, equation (6) implies that

$$v_n^T = \frac{\phi^{(n)}(c_T)}{n!}$$

where

$$\phi(x) = \log(x - c_S)\omega_0^S + \sum_{l=1}^{\infty} \frac{1}{(x - c_S)^l}\omega_l^S$$

Immediately we have

$$v_0^T = \log(c_T - c_S)\omega_0^S + \sum_{l=1}^{\infty} \frac{(-1)^l}{(c_S - c_T)^l}\omega_l^S$$

Differentiating a few times yields

$$\phi'(x) = \frac{1}{x - c_S}\omega_0^S + \sum_{l=1}^{\infty} \frac{-l}{(x - c_S)^{l+1}}\omega_l^S$$

$$\phi''(x) = \frac{-1}{(x - c_S)^2}\omega_0^S + \sum_{l=1}^{\infty} \frac{l(l+1)}{(x - c_S)^{l+2}}\omega_l^S$$

$$\phi'''(x) = \frac{2}{(x - c_S)^3}\omega_0^S + \sum_{l=1}^{\infty} \frac{-l(l+1)(l+2)}{(x - c_S)^{l+3}}\omega_l^S$$

Generalizing gives

$$\phi^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{(x - c_S)^n}\omega_0^S + \sum_{l=1}^{\infty} \frac{(-1)^n}{(x - c_S)^{l+n}} \frac{(l+n-1)!}{(l-1)!}\omega_l^S$$

And from the first equation of this section we get

$$v_n^T = \frac{(-1)^{n-1}}{n(c_T - c_S)^n}\omega_0^S + \sum_{l=1}^{\infty} \frac{(-1)^n}{(c_T - c_S)^{l+n}} \frac{(l+n-1)!}{(l-1)!n!}\omega_l^S$$

$$= -\omega_0^S \frac{1}{n(c_S - c_T)^n} + \sum_{l=1}^{\infty} \frac{(-1)^l}{(c_S - c_T)^{l+n}} \binom{l+n-1}{l-1}\omega_l^S$$

# References


[1] Carrier, J.; Greengard, L.; Rokhlin, V. *A fast adaptive multipole algorithm for particle simulations.* SIAM J. Sci. Statist. Comput. 9 (1988), no. 4, 669–686.

[2] Greengard, L.; Rokhlin, V. *A fast algorithm for particle simulations.* J. Comput. Phys. 73 (1987), no. 2, 325–348.

[3] Greengard, Leslie; Rokhlin, Vladimir *A new version of the fast multipole method for the Laplace equation in three dimensions.* Acta numerica, 1997, 229–269, Acta Numer., 6, Cambridge Univ. Press, Cambridge, 1997.

[4] Greengard L, Lin P (2000) *Spectral approximation of the free-space heat kernel.* Applied and Computational Harmonic Analysis 9:83–97.

[5] Nabors, K.; Korsmeyer, F. T.; Leighton, F. T.; White, J. *Preconditioned, adaptive, multipole-accelerated iterative methods for three-dimensional first-kind integral equations of potential theory.* Iterative methods in numerical linear algebra (Copper Mountain Resort, CO, 1992). SIAM J. Sci. Comput. 15 (1994), no. 3, 713–735.

[6] Beatson, R. and Greengard, L. *A short course on fast multipole methods.* Wavelets, Multilevel Methods and Elliptic PDEs (1997) p. 1-37

[7] Rokhlin, V. *Rapid solution of integral equations of scattering theory in two dimensions.* J. Comput. Phys. 86 (1990), no. 2, 414–439.

[8] Rokhlin, V. *Diagonal forms of translation operators for the Helmholtz equation in three dimensions.* Appl. Comput. Harmon. Anal. 1 (1993), no. 1, 82–93.