Algorithms Fall 2016 Homework 9 Solution

November 22, 2016

Problem 1

(a) The sum of two Toeplitz matrices is Toeplitz. (See part (b).) The product is not necessarily Toeplitz. For example,

$$\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\1 & 1\end{pmatrix} = \begin{pmatrix}2 & 1\\1 & 1\end{pmatrix}$$

(b) For convenience, define $b_i = a_{i,0} = a_{i+1,1}$ for $i = 0,1,...,n-1$. Also define $a_{i,0} = a_{1,1-i}$ for $i = -1,-2,...,-n+1$. Thus we have

$$\left(\begin{array}{cccc|ccc} a_{-3,0} & a_{-2,0} & a_{-1,0} & a_{0,0} & a_{1,0} & a_{2,0} & a_{3,0} \\ \hline & & & & a_{1,1} & a_{2,1} & a_{3,1} \\ & & & & a_{1,2} & a_{1,1} & a_{2,1} \\ & & & & a_{1,3} & a_{1,2} & a_{1,1} \\ & & & & a_{1,4} & a_{1,3} & a_{1,2} \end{array}\right)$$

where the top row (which is not part of the matrix) is also $(a_{1,4}, a_{1,3}, a_{1,2}, a_{1,1} = a_{0,0}, a_{2,1}, a_{3,1}, a_{4,1})$. Observe that $a_{ij} = a_{kl}$ if $j - i = l - k$; this formula extends also to the top (0'th) row. It follows that $a_{j,k} = b_{k-j}$.
It then follows that one can represent a Toeplitz matrix by just the top row of b's ($2n-1$ numbers). The sum of two Toeplitz matrices is represented by the sum of the two corresponding top row, which can be computed in time $O(n)$. (Since any top row leads to a Toeplitx matrix, it follows, in part(a), that the sum of two Toeplitz matrices is Toeplitz).
(c) To multiply by a vector, it is convenient to index the vector backwards: $v = (v_{n-1}, v_{n-2}, ..., v_0)$. Also, we will multiply separately by the upper and lower triangle:

$$\begin{pmatrix}b_0 & b_1 & b_2 & b_3\\0 & b_0 & b_1 & b_2\\0 & 0 & b_0 & b_1\\0 & 0 & 0 & b_0\end{pmatrix}\cdot\begin{pmatrix}v_{n-1}\\v_{n-1}\\.\\.\\.\\v_1\\v_0\end{pmatrix}$$

The resulting vector product is n of the $2n-1$ terms in the convolution of the $b_{i\geq 0}$ sequence and the v sequence. We can do this in time $O(n\log n)$. Similarly, we can multiply the lower triangle by v by convolving the $b_{i<0}$ sequence with the v sequence. We then add the two vectors of length n.
(d) We can multiply a Toeplitz matrix by an arbitrary matrix M in time $n^2 \log n$ by multiplying by each column of M separately. Some speedups are possible if M is also Toeplitz, but note that we need to output n^2 numbers.
ADDITIONAL COMMENTARY: Is somewhat more natural and elegant to define the convolution as $(f \otimes g)(k) = \sum_{0\leq j<n} f_i g_{k-j}$, where k and $k-j$ are taken modulo n. Thus there are only n elements in $f \otimes g$, not $2n-1$. This corresponds to multiplying polynomials modulo $x^n - 1$, rather than multiplying polynomials without modular reduction. This may be regarded as an alternative to padding with zeros.
We then have the formula that the Fourier transform of $(f \otimes g)$ is the pointwise product of the Fourier transform of f and the Fourier transform of g.

This corresponds to the circulant variation of Toeplitz matrices, of the form

$$\begin{pmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{pmatrix}$$

Note that a circulant matrix is a special kind of Toeplitz matrix. The sum of two circulant matrices is circulant and the product of two circulants is circulant, which can be checked easily. One can represent a circulant matrix by its top row (which is now part of the matrix). To multiply two circulants, take the Fourier transforms of each top row, multiply those together, then form a circulant from the result. This takes time $O(n \log n)$.