Inverse of the Square Wave Matrix

David J. Grabiner
then a grad. student at Math. dept. Harvard University
Cambridge, MA 02138

Warren D. Smith
then at NEC Research Institute
4 Independence Way, Princeton, NJ 08540;
now 21 Shore Oaks Drive, Stony Brook NY 11790;
WDSmith@fastmail.fm

Sushanta Srivastava
then at AT&T Bell Laboratories
East Windsor, NJ 08512

January 27, 2005

Consider the $N \times N$ matrix A, whose ijth element $A_{i,j}$, $1 \leq i, j \leq N$, is defined by

$$A_{i,j} = (-1)^{\lceil j/i \rceil + 1}. \tag{1}$$

The ith row of A represents a ± 1-valued "square wave" function of j with half-period i. Thus $\mathbf{x}A$, where $\mathbf{x}$ is a (row) N-vector, gives a picture of the waveform that is a linear combination of square waves with weights in $\mathbf{x}$. Similarly $\mathbf{y}A^{-1}$ finds weights so that a given waveform $\mathbf{y}$ is a weighted linear combination of these square waves.

This matrix has a surprisingly simple inverse. More generally, we can invert any matrix in which the first row is arbitrary, except that it cannot begin with 0, and the ith row repeats each term in the first row i times until it reaches the end of the row, thus spreading out the function with larger periods. The matrix A will thus have entries

$$A_{i,j} = f\left(\left\lceil \frac{j}{i} \right\rceil\right). \tag{2}$$

Let $g(n)$ be the *Dirichlet inverse* [1, p. 30] of the function $f(n) - f(n+1)$, that is, the function which satisfies

$$\sum_{k|n} g\left(\frac{n}{k}\right)[f(k) - f(k+1)] = \delta_{n,1}, \tag{3}$$

where $\delta_{i,j}$ is Kronecker's delta function, and k dividing n is denoted by $k \mid n$. Extend the domain of $g(x)$ to $x \in \overline{\mathbf{R}}$ by defining $g(x) = 0$ for $x \notin \mathbf{Z}^+$, including $x = \infty$.

Theorem. *The inverse of the matrix A is the matrix B given by*

$$B_{i,j} = \begin{cases} g\left(\frac{j}{i}\right) - g\left(\frac{j}{i-1}\right), & \text{if } 1 \le j < N; \\ \frac{1}{f(1)}\delta_{i,1} - \sum\limits_{1 \le k < N} B_{i,k}, & \text{if } j = N. \end{cases} \tag{4}$$

Remark. This theorem shows that A^{-1} is a sparse matrix. Since $B_{i,j}$ can only be nonzero if $j = N$, or if j/i or $j/(i-1)$ is an integer, the number of nonzero entries in B (or in A^{-1}) is at most $2N \ln N + O(N)$.

Proof. Let $AB = C$ and consider $C_{i,k}$. We wish to show $C_{i,k} = \delta_{i,k}$.

Every element in the nth row of A is $f(1)$. Thus, if $1 \le k < i = N$, then we have

$$C_{N,k} = f(1) \sum_{1 \le j \le N} \left[g\left(\frac{k}{j}\right) - g\left(\frac{k}{j-1}\right) \right] = 0 \tag{5}$$

as required, by telescopic cancellation. (Since $g(k/0) = g(k/N) = 0$, no problems can arise at the ends of the summation.) Similarly, if $i = k = N$ we have $C_{N,N} = 1$ by summing (5) along rows.

Therefore we see that proving (4) is equivalent to showing the identity

$$\sum_{1 \le j \le N} f(\lceil j/i \rceil) \left[g\left(\frac{k}{j}\right) - g\left(\frac{k}{j-1}\right) \right] = \delta_{i,k} \tag{6}$$

if $1 \le i, k < N$. (Once this is proven, $C_{i,N} = 0$ will follow from summing (6) along rows.)

The only terms in (6) which do not telescopically cancel are the $g(k/j)$ with i dividing j. Since $g(k/j) = 0$ unless $j \mid k$, the left hand side of (6) is

$$\sum_{\substack{1 \le j < N \\ i|j|k}} \left[f\left(\frac{j}{i}\right) - f\left(1 + \frac{j}{i}\right) \right] g\left(\frac{k}{j}\right). \tag{7}$$

and this is $\delta_{k/i,1} = \delta_{k,i}$ by (3) with $n = k/i$. □

The usefulness of this theorem is exemplified in the following corollary.

Corollary. *Multiplication of an N-vector by either A or $B = A^{-1}$ can be accomplished in $O(N \log N)$ operations, once the matrix B is known.*

Proof. Since B has only $O(N \log N)$ nonzero entries, it can be multiplied by an N-vector in $O(N \log N)$ operations. Multiplying an N-vector by A may also be accomplished in $O(N \log N)$ operations, since B is upper Hessenberg ($B_{i,j} = 0$ if $i > j + 1$). More precisely: If the first element of $\mathbf{y}$ were known, then computing the remaining elements of $\mathbf{y} = \mathbf{x}A = \mathbf{x}B^{-1}$ from the first $N-1$ elements of $\mathbf{x}$ would be a sparse back substitution problem, since the minor obtained from B by deleting row 1 and column N is upper triangular. This first element may be found by using only the first column of A; from (2), we have $y_1 = f(1)\sum_{i=1}^{N} x_i$. $\square$

Thus, if the first N values of g can be computed in $O(N \log N)$ steps, then a vector can be multiplied by either A or A^{-1} in $O(N \log N)$ steps. Once they have been computed, regardless of the complexity of this computation, any vector of length N or less can be multiplied by the appropriately-sized matrix A or B in $O(N \log N)$ steps.

Now, consider the special case in which $f(n) = (-1)^{n+1}$, giving the square wave matrix (1).

Lemma. *The inverse function $g(n)$ in (3) in this case is given as follows. If n has the prime factorization*

$$n = 2^r p_1^{e_1} p_2^{e_2} \cdots p_\omega^{e_\omega} \tag{8}$$

where the p_i are odd primes, then $g(n) = Q(n)/2$, where

$$Q(n) = \begin{cases} 0, & \text{if } \max_i e_i \geq 2; \\ (-1)^\omega 2^{\max(r-1,0)}, & \text{otherwise.} \end{cases} \tag{9}$$

Proof. Note that $f(n) - f(n+1) = 2f(n)$, so we need to show that $f(n)$ and $Q(n)$ are Dirichlet inverses, that is,

$$\sum_{d|n} (-1)^{d+1} Q(n/d) = \delta_{n,1}. \tag{10}$$

Let oddpart(n) denote what is left after n is repeatedly divided by 2 until it becomes odd, and let $d = 2^a b$ where $b = \operatorname{oddpart}(d)$. We may now rewrite (10) as

$$-\sum_a \sum_{b\ \text{odd}, b|n} (-1)^{2^a b} Q\left(\frac{n}{2^a b}\right) \tag{11}$$

Define m by $n = 2^m \operatorname{oddpart}(n)$. We now use the fact that, for positive integer $k = 2^r \operatorname{oddpart}(k)$,

$$Q(k) = 2^{\max(r-1,0)} \mu(\operatorname{oddpart}(k)), \tag{12}$$

where $\mu(n)$ is the Möbius function from number theory [1], to rewrite (11) as

$$-\sum_a \sum_{b\ \text{odd}, b|n} (-1)^{2^a b} 2^{\max(m-a-1,0)} \mu\left(\operatorname{oddpart}\left(\frac{n}{2^a b}\right)\right). \tag{13}$$

The inner sum may be evaluated by using the simplest form of the Möbius inversion formula

$$\sum_{d|n} \mu(d) = \delta_{n,1}, \tag{14}$$

valid for all positive integers n. Since $(-1)^{2^a b} = (-1)^{2^a}$, the result is

$$-\sum_{0\le a\le m} (-1)^{2^a} 2^{\max(m-a-1,0)} \delta_{\text{oddpart}(n),1}, \tag{15}$$

and now using the geometric sum

$$2^{\max(m-1,0)} - \sum_{1\le a\le m} 2^{m-a-1} = \delta_{m,0} \tag{16}$$

(for $m \ge 0$) yields

$$2\delta_{\text{oddpart}(n),1}\delta_{m,0}. \tag{17}$$

But if $m = 0$, then n is odd, so that oddpart$(n) = n$. Hence this is $2\delta_{n,1}$, and the proof is complete. □

This gives us a formula for A^{-1}. If A is the square wave matrix (1), then $B' = 2A^{-1}$ is given by

$$B'_{i,j} = \begin{cases} Q\left(\frac{j}{i}\right) - Q\left(\frac{j}{i-1}\right), & \text{if } 1 \le j < N; \\ 2\delta_{i,1} - \sum\limits_{1\le k<N} B'_{i,k}, & \text{if j=N.} \end{cases} \tag{18}$$

Square waves are readily generated by digital circuitry, and the response of linear circuits to square waves is readily predicted. It is computationally feasible to resolve a signal into square waves, because multiplication of an N-vector by A or A^{-1} can be accomplished in $O(N \log N)$ shift and add operations only. Every entry of B', except in the last column, is a sum or difference of at most two powers of 2. Thus every multiplication by these numbers takes at most two shifts and adds. The last column of B' can be handled by using the summation in (4), which requires an additional $O(N \log N)$ adds. In the back substitution process, no divisions are required because the subdiagonal entries of B', which become the divisors during the back substitution, are all -1; the multiplications can again be done with $O(N \log N)$ shifts and adds.

Even if bit shifting takes time proportional to the distance shifted, we still need only $O(N \log N)$ time, because the bit shifts due to $Q\left(\frac{mi}{i}\right)$, excluding the terms in the last column, occur only in the terms $A_{i,mi}$, $A_{i,n}$, $A_{i+1,mi}$, and $A_{i+1,n}$. Thus the total length of all the bit shifts for a fixed i is

$$2\sum_{m=1}^{\lfloor \frac{N}{i} \rfloor} \max(0, (\text{exponent of 2 in factorization of } m) - 1)$$

$$
\begin{aligned}
&= 2\left(\left\lfloor\frac{N}{4i}\right\rfloor + \left\lfloor\frac{N}{8i}\right\rfloor + \left\lfloor\frac{N}{16i}\right\rfloor + \cdots\right) \\
&< \frac{N}{i},
\end{aligned}
$$

and the total length for all i is thus $N \ln N + O(N)$. This is all the shifts that are needed, since the last column can be handled by summation.

By using column operations to make A upper triangular, one easily verifies that $\det(A) = 2^{N-1}$ for the square wave matrix.

Another interesting case is $f(n) = n$, which gives the matrix $A_{i,j} = \lceil j/i \rceil$. Here, $f(n) - f(n+1) = -1$, so we get $g(n) = -\mu(n)$. Thus the inverse matrix B is given by

$$
B_{i,j} = \begin{cases} \mu\left(\frac{j}{i-1}\right) - \mu\left(\frac{j}{i}\right), & \text{if } 1 \le j < N; \\ \delta_{i,1} - \sum\limits_{1 \le k < N} B_{i,k}, & \text{if } j = N. \end{cases} \tag{19}
$$

The third author [2] discovered the inverse of the square wave matrix several years ago while investigating signal representation. He also invented the function $Q(x)$ and used several of its properties to prove the identity $AB' = 2I$. The compact proof presented here is largely a contribution of the first two authors.

References

[1] Apostol, Thomas M., *Introduction to analytic number theory,* Springer-Verlag 1976.

[2] Srivastava, Sushanta, *Analysis of Digital Signals by a Set of Equitransition Binary Functions,* unpublished manuscript, November 1984.