Procedure

Rohan Mukherjee

July 24, 2024

We start by defining the following heuristic.

procedure FindMatrix(m, n)
  Initialize a $2^{n-1} \times n$ matrix A whose rows are the vectors $\{1\} \times \{\pm 1\}^{n-1}$
  while number of rows in $A > m$ **do**
    $a \leftarrow$ any (there may be many) row of A with the smallest number of zero entries
    if A has a row b differing in only one entry i from a with $b_i = \pm 1$ and $a_i = \mp 1$ **then**
      Replace rows a, b of A with one row equal to $\frac{a+b}{2}$ (equiv. zero out entry i)
    else
      The procedure fails.
    end if
  end while
  return A after normalizing its rows
end procedure

The else condition is very important: this procedure cannot always be carried out. We illustrate this by giving a 5×4 matrix that you cannot make into a 4×4 matrix.

Consider the following instance of the procedure, where we have highlighted rows of the same color that get combined:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix} \tag{1}$$

Some of these vectors have 1 zero, and others have no zeros. By the procedure, we would not be allowed to combine vectors with 1 zero until we have combined all vectors that have no zeros. The only rows with no zeros are $(1,1,1,1)$ and $(1,-1,-1,-1)$. So such a b doesn't exist for the vector $(1,1,1,1)$, since the only possible option, $(1,-1,-1,-1)$, differs in 3 entries rather than just 1.

We can use this procedure to give many different examples. There is a natural rule that comes to mind. Let $\preceq$ be the lexicorigraphical order on $\mathbb{R}^n$. Choose a according to the rule that a is maximal w.r.t. $\preceq$ among A's rows with the fewest number of zero entries. Then always pick b to be the unique vector that differs in only the last nonzero entry of a. For a 4×4 example, one gets:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix}$$

$$\mapsto \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & -1 & 0 \end{bmatrix}$$

The best part about the procedure is that it works for matrices of any size. Indeed, we could've stopped at any of the above steps to get matrices of size 8×4, 7×4, 6×4, and 5×4 as well.

If one carries out the procedure using the rules in the first example, one gets the following:

Theorem 1. *Let A be the $n \times n$ matrix that results from the procedure by always choosing a to be the unique vector that is maximal w.r.t. the lexicographical order $\preceq$ among the rows of A with the fewest number of zeros, there exists a unique b that differs in only the last nonzero entry of a. Then,*

$$\beta(A) = 2\sqrt{\lfloor \log_2(n) \rfloor + 1} - \sqrt{\lfloor \log_2(n) \rfloor + 2} + \frac{n}{2^{\lfloor \log_2(n) \rfloor}} \left(\sqrt{\lfloor \log_2(n) \rfloor + 2} - \sqrt{\lfloor \log_2(n) \rfloor + 1} \right)$$

Proof. Fix $k = \lfloor \log_2(n) \rfloor$, and assume that n is not a power of 2, so that $2^k < n < 2^{k+1}$.

We highlight a brief sketch of the proof here. We first claim that the $n \times n$ matrix this method generates will be the same as a $n \times n$ matrix with a $n \times (k+2)$ block, with the rest of the columkns being padded 0s. Padding columns of zeros then ends up giving the exact same β value.

Then to find the β value, we will notice that the initial matrix has rows $\{1\} \times \{\pm 1\}^{k+1}$. It will turn out that, after normalizing, having $k+1$ nonzero entries will yield a value of $\sqrt{k+1}$ on an entire group (the W_i as in the structure theorem), by the choice of the groups.

Finally, we can count the number of rows that were combined and not combined: these turn out to be $2^{k+1} - n$ and $n - (2^{k+1} - n) = 2n - 2^{k+1}$ respectively. The rows that were combined will have only $k+1$ nonzero entries, while the rows that are not combined end up having $k+2$ nonzero entries. This will yield the β value from above.

Now, let $A \in \mathbb{R}^{m \times n}$ be a matrix with normalized rows. We claim that if $A' = [A\ 0]$ where $0 \in \mathbb{R}^m$ is the all 0s vector, then $\beta(A') = \beta(A)$. Writing $x' = [x, \pm 1]$ for $x \in \{\pm 1\}^n$, we see that:

$$\beta(A') = \frac{1}{2^{n+1}} \sum_{x' = [x, \pm 1] \in \{\pm 1\}^{n+1}} \|A'x\|_\infty = \frac{2}{2^{n+1}} \sum_{x \in \{\pm 1\}^n} \|Ax\|_\infty = \beta(A)$$

Since $Ax = A'[x, \pm 1]$. Then it suffices to just ignore columns of padded 0s.

We let B be the $2^{n-1} \times n$ matrix where the rows are $\{1\} \times \{\pm 1\}^{n-1}$ ordered lexicorigraphically. Then in each step of the procedure described in the theorem, we are just combining the topmost row with the fewest amount of zeros with the row directly below it. This is because the row directly below it will only differ in the last nonzero entry having a minus sign instead (by the lexicorigraphical ordering).

Thus by construction of the procedure, since we combine two rows only if they differ in their last nonzero entry, zeros only show up in the last columns. Repeating this process, instead of starting with the $2^{n-1} \times n$ matrix and averaging all the way down, we can instead

start with the $2^{k+1} \times (k+2)$ matrix whose rows are all the vectors in $\{1\} \times \{\pm 1\}^{k+1}$ ordered lexicorigraphically and continue averaging rows until we have only n rows (just remove the trailling columns of zeros once the procedure has hit 2^{k+1} rows).

We say that a row a is useful to a hypercube vertex x if a maximizes $|a^\top x|$ among all rows of A. Then notice that for this very tall matrix, each row has precisely two hypercube vertices that it is useful to: the row itself and it's negative after ignoring the first column. For example in the $2^3 \times 4$ matrix from before the row $[1,1,1,1]$ is useful to only $[1,1,1,1]$ and $[-1,-1,-1,-1]$ (that is to say, for every other hypercube vertex, there is another row of a that makes the dot product larger). This is saying that a vector of the hypercube will be useful to these rows whose last entries are all zeros iff the prefix of the hypercube vertex is the same as the row (up to negating the hypercube vector). As another example, $[1,1,0,0]$ would be useful to $[1,1,\pm 1,\pm 1]$ and $[-1,-1,\pm 1,\pm 1]$. In the case where we have removed all the columns of zeros, there can only ever be two vertices that are useful to a row, since rows can only have at most 1 zero.

We ignore the "up to negating" factor of two by doing the following. Since $\{1\} \times \{\pm 1\}^{n-1} = -\{-1\} \times \{\pm 1\}^{n-1}$, for any matrix A we have that:

$$\beta(A) = \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} \|Ax\|_\infty = \frac{1}{2^{n-1}} \sum_{x \in \{1\} \times \{\pm 1\}^{n-1}} \|Ax\|_\infty$$

Now we can simply count the number of rows that we need to combine: we start with 2^{k+1} rows and we need to have only n rows. Thus we need to remove $2^{k+1} - n$ rows, and notice that averaging two rows into one removes precisely one row. Thus in the final matrix $2^{k+1} - n$ are rows that are the average of two others rows, and the rest are just hypercube vertices on their own.

Also notice that averaging two rows that differ in only the last coordinate simply makes this last coordinate 0. In this way, the vertices that are useful to this new averaged row is simply the union of the vertices that were useful to each row respectively. For example, when we averaged rows $[1,1,1,1]$ with row $[1,1,1,-1]$, we got the row $[1,1,1,0]$ which is useful to precisely the set of vertices who start with $[1,1,1,0]$ of which there are precisely two.

So for rows that were averaged, there are precisely 2 vertices in their group, being each of the rows themselves, and if the row was not averaged, then it's only useful to itself. Now we can count the β value on these rows. If $a \in \mathbb{R}^m$ is a row whose first ℓ entries are either 1 or -1, then for every vertex of the hypercube $x \in \{\pm 1\}^m$ with the same prefix, i.e. agreeing with a in those first ℓ entries, we see that $|a^\top x| = \ell$. After normalizing a since we

require normalized rows, we get that $|a^\top x| = \sqrt{\ell}$ for every vertex that agrees with it in all of it's non-zero entries.

Now we move to the case of combined rows. By the above, in this simplified case where we have artificially removed all the columns of 0s, we only combine two rows if they differ only in the last entry. Since these rows initially have $k+2$ nonzero entries, after averaging they have $k+1$ nonzero entries. By the calculation from before, after normalizing they give a value of $\sqrt{k+1}$ on their entire group, which is of size 2.

Similarly, if a row was not combined, then it has all $\sqrt{k+2}$ nonzero entries and it's group has size 1. Now we are almost done. We simply have to count then number of combined rows, and we can use these last two facts to find the β value. Since we start with 2^{k+1} rows and we want n rows, we combine 2 rows $2^{k+1} - n$ times as discussed above. Thus in the final $n \times k+2$ matrix we have $2^{k+1} - n$ combined rows. The remaining $n - (2^{k+1} - n) = 2n - 2^{k+1}$ rows are not combined. Using the above observations, this gives a beta value of:

$$\frac{\sqrt{k+1} \cdot 2 \cdot (2^{k+1} - n) + \sqrt{k+2} \cdot 1 \cdot \left(2n - 2^{k+1}\right)}{2^{k+1}}$$

After some algebra, we get:

$$2\sqrt{k+1} - \sqrt{k+2} + \frac{n}{2^k}\left(\sqrt{k+2} - \sqrt{k+1}\right)$$

The simpler case of when $n = 2^k$ is a power of 2 will yield the matrix:

$$\begin{bmatrix} 1 & B_k & 0 \end{bmatrix}$$

Where B_k is the $2^k \times k$ matrix of all ± 1 combinations as in the proof of Theorem 9. From here we see similarly that if we ignore the columns of 0s, we only need to find the β value of the below matrix:

$$\begin{bmatrix} 1 & B_k \end{bmatrix}$$

Theorem 9 tells us that the β value for this is just $\sqrt{k+1}$. This completes the proof. □