Almost Orthogonal Vectors

Yajit Jain, Deepak Narayanan, Leon Zhang

December 10, 2014

1 Introduction

Consider a collection of N unit vectors $v_1, \ldots, v_N$ in $\mathbb{R}^n$. Define ϵ as

$$\epsilon = \max_{i \neq j} |v_i \cdot v_j|^2.$$

In this paper, we try to address the question of what the minimal attainable ϵ is.

For $n = N$, it is clear that the optimal value of ϵ is zero, obtained by letting $v_1, \ldots, v_N$ be any orthonormal basis. And for $N > n$, we must have $\epsilon > 0$ as any collection of N vectors in n-space is linearly dependent, while any collection of orthonormal vectors is linearly independent.

For fixed n, it is also easy to see that the minimum value of ϵ is nondecreasing as N increases: that is, if ϵ_1 is the minimum value of ϵ for N unit vectors in $\mathbb{R}^n$, and ϵ_2 is the minimum value of ϵ for $N+1$ unit vectors in $\mathbb{R}^n$, we must have

$$\epsilon_1 \leq \epsilon_2.$$

Otherwise, given a configuration of $N+1$ vectors in $\mathbb{R}^n$ with $\epsilon = \epsilon_2$, we could simply remove one vector and obtain a configuration of N vectors whose ϵ would be less than or equal to ϵ_1.

Along the same lines, given a configuration of N vectors in n-space with corresponding ϵ, we can obtain a collection of $N+1$ vectors in $\mathbb{R}^{n+1}$ with the same ϵ: simply embed the first N vectors in $\mathbb{R}^n$, then add any vector perpendicular to the hyperplane in which the N vectors lie. Hence the minimum ϵ for collections of $N+1$ vectors in $\mathbb{R}^{n+1}$ is upper-bounded by the corresponding ϵ for N vectors in $\mathbb{R}^n$.

In Section 2, we shall see that for $n = 2$, the minimum value of ϵ is given by

$$\epsilon = \cos^2 \pi/N.$$

We will then derive in Section 3 a lower bound on ϵ for any choice of n and N. In Section 4 we shall see that, for $N = n+1$, the minimum value of ϵ is given by

$$\epsilon = \frac{1}{n^2},$$

and this configuration is produced when the $n+1$ vectors point to the vertices of the n-simplex. We will also exhibit an optimal configuration of six vectors in $\mathbb{R}^3$; this optimal configuration produces an ϵ equal to

$$\epsilon = \frac{1}{5}.$$

In Section 5, we present the results of a computational approach to computing minimal values of ϵ. Finally, in Section 6 we consider the related problem of minimizing

$$\delta = \max_{i \neq j} v_i \cdot v_j.$$

In particular, we will show that in the case where $N = n+1$, the simplex remains the optimal configuration.

Leon wrote this section.

2 N vectors in $\mathbb{R}^2$

In this section, we consider the problem of almost orthogonal vectors in $\mathbb{R}^2$.

As stated previously, when $N = 2$ we can easily obtain an ϵ equal to 0; a simple example of such a configuration is $[1,0]^T$ and $[0,1]^T$.

The problem, however, gets harder for larger N. Let us first consider the specific case of $N = 3$. Once we build some intuition for the problem, we will generalize the result for arbitrary N.

Theorem 2.1. *For $n = 2$ and $N = 3$, the minimum possible ϵ is equal to $1/4$.*

Before proving the above theorem, we state and prove the following lemma, which will be useful for the general case as well.

Lemma 2.2. *Consider n angles $\theta_1, \theta_2, \ldots, \theta_n \in [0, \pi]$ s.t. $\theta_1 + \theta_2 + \ldots + \theta_n = \pi$. Then $\alpha = \max_i \cos^2 \theta_i$ must equal $\cos^2 \theta_j$ where $j = \text{argmin } \theta_i$.*

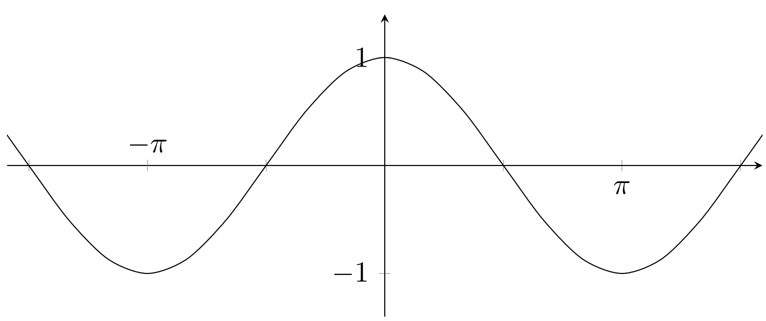

Figure 1: Plot of $\cos\theta$ versus θ

Proof. Our argument hinges on the fact that for all $\theta \in [0, \pi/2]$, the function $\cos\theta$ is decreasing – this is easy to see from Figure 1.

Without loss of generality, let us assume that $\theta_1 \geq \theta_2 \geq \ldots \geq \theta_n$. Hence our theorem statement is equivalent to proving that α is equal to $\cos^2\theta_n$.

We split our proof into two cases,

- $\theta_1, \theta_2, \ldots, \theta_n \in [0, \pi/2]$: In this case, it is easy to see that $\alpha = \cos^2\theta_n$ from the fact that $\cos\theta$ is a decreasing function in θ if $\theta \in [0, \pi/2]$.

- One of $\theta_1, \theta_2, \ldots, \theta_n$ is greater than $\pi/2$: Then in particular $\theta_1 > \pi/2$. We see that $\cos^2\theta_1 = \cos^2(\pi - \theta_1)$ which is equal to $\cos^2(\theta_2 + \theta_3 + \ldots + \theta_n)$. Furthermore, $\pi - \theta_1 = \theta_2 + \theta_3 + \ldots + \theta_n < \pi/2$, which means $\cos^2\theta_1 = \cos^2(\theta_2 + \theta_3 + \ldots + \theta_n) < \cos^2\theta_n$. (since $\theta_n < \theta_2 + \theta_3 + \ldots + \theta_n$)

  In addition, we see that for all $i \in \{2, 3, \ldots, n-1\}$, $\cos^2\theta_i < \cos^2\theta_n$, from which we can conclude that α is still equal to $\cos^2\theta_n$.

□

Given this lemma, we now prove the theorem stated above.

Proof. Observe that for $N = 3$, the quantity ϵ for three unit-length vectors $v_1, v_2, v_3 \in \mathbb{R}^2$ is given by

$$\max\{|v_1 \cdot v_2|^2, |v_1 \cdot v_3|^2, |v_2 \cdot v_3|^2\}$$

Because the sign of the dot product of any two vectors does not matter, and because

$$|v \cdot v'| = |v \cdot (-v')|$$

for some arbitrary vector v', we see that it's easy to transform the three vectors v_1, v_2 and v_3 so that they all lie in the same semi-circle – to accomplish this, at most one vector needs to be reflected about the origin (multiplied by -1).

Without loss of generality, let $v_1 = [1, 0]^T$. Also, without loss of generality let us assume that v_2 and v_3 are above the x-axis, and that v_1, v_2 and v_3 are in anti-clockwise order.

Let θ_1 be the angle between v_1 and v_2 and θ_2 be the angle between v_2 and v_3. Let us now define θ_3 such that $\theta_1 + \theta_2 + \theta_3 = \pi$. Observe that $|v_1 \cdot v_3|^2 = \cos^2(\theta_1 + \theta_2) = \cos^2(\pi - \theta_3) = \cos^2\theta_3$, hence we can conclude that

$$\epsilon = \max_{i \in \{1,2,3\}} \cos^2\theta_i$$

From the above lemma, if $j = \text{argmin}\ \theta_i$ and $\theta_1 + \theta_2 + \theta_3 = \pi$, then $\epsilon = \cos^2\theta_j$.

Furthermore, we can obtain an upper bound on θ_j by observing that $\theta_1 + \theta_2 + \theta_3 \geq \theta_j + \theta_j + \theta_j = 3\theta_j \Rightarrow \theta_j \leq \pi/3$. Since we're interested in the smallest such ϵ and since the cosine function is a decreasing function in θ between 0 and $\pi/2$, the optimum value of ϵ for $n = 2$ and $N = 3$ is $\cos^2\pi/3$. Figure 2 shows this optimum configuration.

□

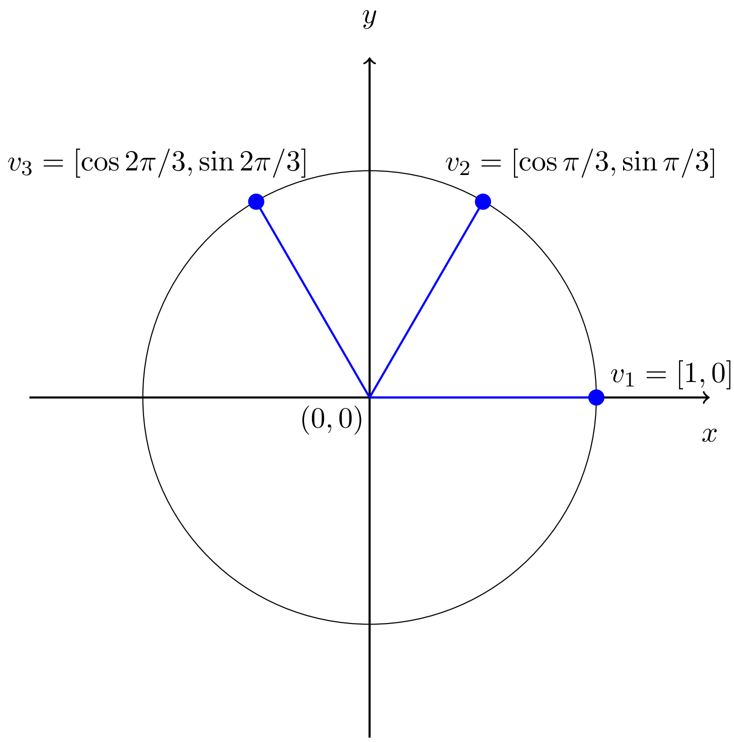

Figure 2: A configuration of three unit vectors that produce the optimum ϵ for $n = 2, N = 3$

Given the above result for $N = 3$, we attempt to generalize to any integer N in the following theorem.

Theorem 2.3. *For $n = 2$ and arbitrary N, the optimum value of ϵ is given by $\cos^2 \pi/N$.*

Proof. The proof of this theorem is similar to the proof for the specific case of $N = 3$.

Again without loss of generality, we can assume that the vectors $v_1, v_2, \ldots, v_N$ are on or above the x-axis, and that $v_1 = [1, 0]^T$ – if any vector v_i were not above the x-axis, then we could just consider $-v_i$ (the reflection of v_i about the origin) instead.

Let us define θ_i as the angle between the vectors v_i and v_{i+1} for $i \in \{1, 2, \ldots, N-1\}$, and let θ_N be the angle such that $\theta_1 + \theta_2 + \ldots + \theta_N = \pi$.

Observe that as before, we are interested in maximizing the square of the cosine of the angle between any two vectors in $v_1, v_2, \ldots, v_N$. Note that here, the angle between any two vectors in v_j and v_k in $v_1, v_2, \ldots, v_N$ can be expressed as $\sum_{i=j}^{k-1} \theta_i$. Note that, however, if $\sum_{i=j}^{k-1} \theta_i \leq \pi/2$, then $\cos^2 \theta_j \geq \cos^2(\sum_{i=j}^{k-1} \theta_i)$ and that if $\sum_{i=j}^{k-1} \theta_i > \pi/2$, then $\cos^2(\sum_{i=j}^{k-1} \theta_i) = \cos^2(\pi - \sum_{i=j}^{k-1} \theta_i)$ which is less than $\cos^2 \theta_i$ for any i in

$$\{1, 2, \ldots, N-1, N\} \setminus \{j, j+1, \ldots, k-1\},$$

where $j, k \in \{1, 2, \ldots, N-1\}$ and $j \neq k$. (the above set must be non-empty since it must at least contain the element N)

From this we can conclude that ϵ is equal to $\max_i \cos^2 \theta_i$. Then, if $j = \text{argmin } \theta_i$, we see that $\epsilon = \cos^2 \theta_j$ by Lemma 2.2.

Since $\theta_1 + \theta_2 + \ldots + \theta_N \geq N \cdot \theta_j \Rightarrow \theta_j \leq \pi/N$, we see that the optimum ϵ value is in fact equal to $\cos^2 \pi/N$ (again making use of the fact that the cosine function is decreasing between 0 and $\pi/2$). □

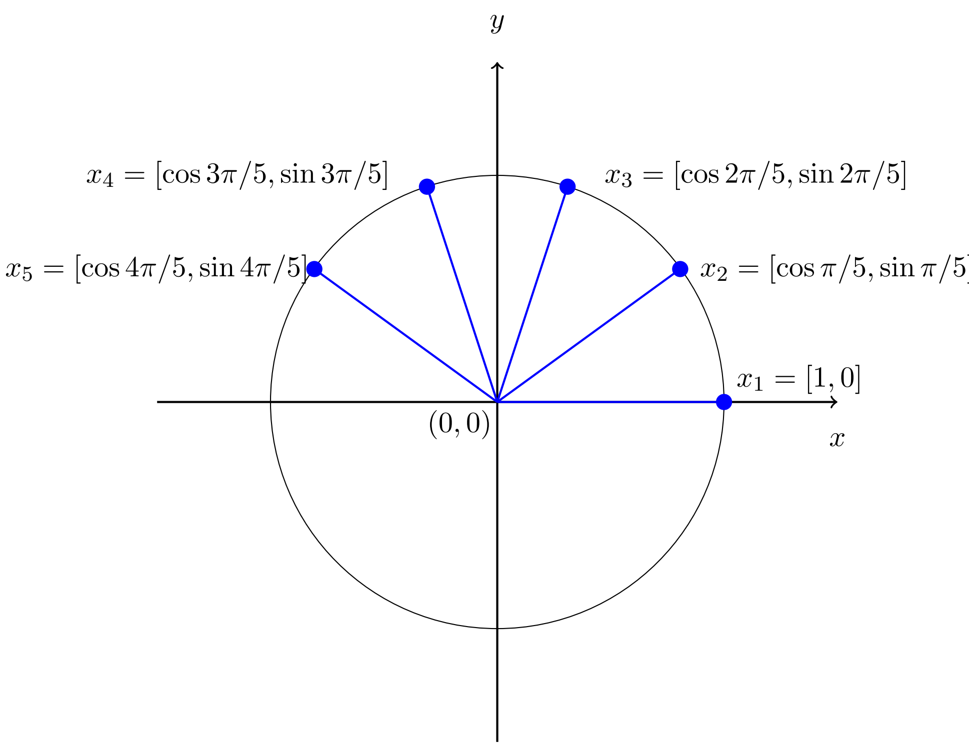

Figure 3: A configuration of five unit vectors that produce the optimum ϵ for $n = 2, N = 5$

Figure 3 shows an optimum configuration for $n = 2$ and $N = 5$. Note by taking reflections of x_2 and x_4 in Figure 3, we get five vectors that form a regular 5-gon – this can be seen in Figure 4. A similar transformation can be done for all odd N.

Deepak wrote this section. Leon edited this section. Yajit made minor edits to this section.

3 A lower bound on ϵ

In the search for a lower bound on ϵ for arbitrary (n, N) pairs, it is natural to consider the so-called *Gram matrix* of our collection of vectors.

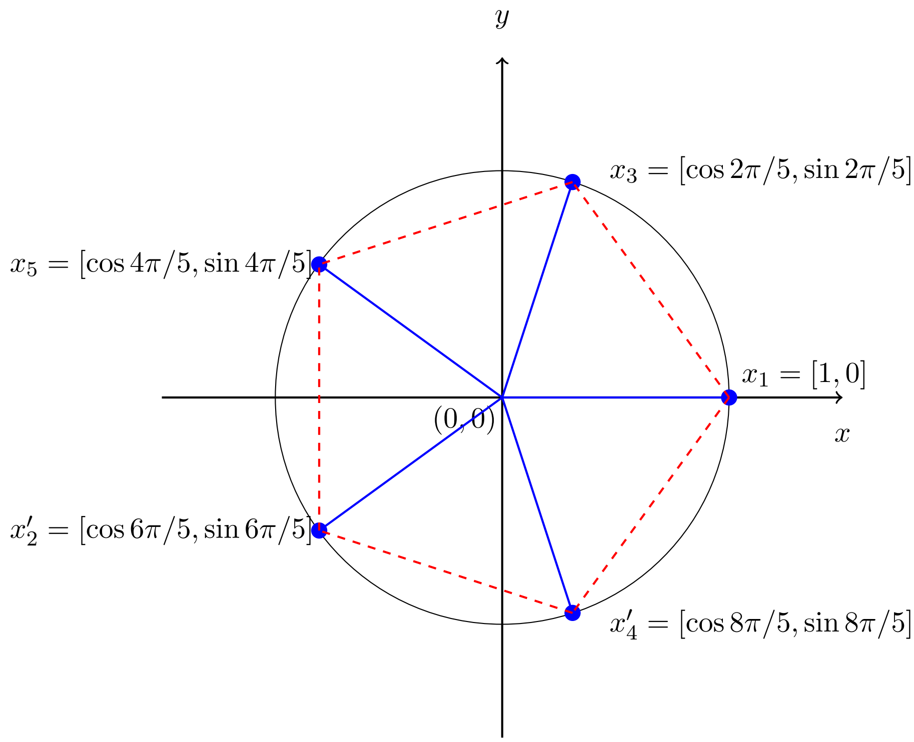

Figure 4: A configuration of five unit vectors that produce the optimum ϵ for $n = 2, N = 5$ with x_2 and x_4 from the previous figure reversed

---

Definition 3.1. *The Gram matrix of a collection of N vectors $v_1, \ldots, v_N$ in $\mathbb{R}^n$ is the N-by-N matrix given by*

$$G(v_1, \ldots, v_N) = \begin{pmatrix} v_1 \cdot v_1 & v_1 \cdot v_2 & \ldots & v_1 \cdot v_n \\ v_2 \cdot v_2 & v_2 \cdot v_2 & \ldots & v_2 \cdot v_n \\ \vdots & \vdots & \ddots & \vdots \\ v_n \cdot v_1 & v_n \cdot v_2 & \ldots & v_n \cdot v_n \end{pmatrix}.$$

Note that the Gram matrix can be written as the product $A^T A$, where A is the n-by-N matrix

$$A = \begin{pmatrix} | & | & \cdots & | & | \\ v_1 & v_2 & \cdots & v_{N-1} & v_N \\ | & | & \cdots & | & | \end{pmatrix}.$$

We can use a well-known lemma on the rank of certain real symmetric matrices to derive a useful lower bound on ϵ. Our proof of the lemma follows that of a paper by Noga Alon[1].

[1] Lemma 2.2, *Perturbed identity matrices have high rank: proof and applications*, Noga Alon.

Lemma 3.2. *Let $A = (a_{i,j})$ be a n-by-n real symmetric matrix with $a_{i,i} = 1$ for all i and $|a_{i,j}| \leq \sqrt{\epsilon}$ for all $i \neq j$. Let d be the rank of A. Then*

$$\epsilon \geq \frac{n-d}{d(n-1)}.$$

Proof. We define $\lambda_1, \ldots, \lambda_n$ to be the eigenvalues of A; write k as the number of nonzero eigenvalues. Their sum is the trace of A, equal to n. We know that the number of nonzero eigenvalues of a complex matrix is less than or equal to the rank of the matrix: so we have that $d \geq k$.

Recall the Cauchy-Schwarz inequality for $\mathbb{R}^n$: given $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n) \in \mathbb{R}^n$, we have

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right).$$

In our case, let the x_i equal λ_i, and let each y_i equal 1 when λ_i is nonzero and zero when λ_i is zero. We get

$$n^2 = \left(\sum_{i=1}^{n} \lambda_i\right)^2 \leq \left(\sum_{i=1}^{n} \lambda_i^2\right) k$$

so that

$$\sum_{i=1}^{n} \lambda_i^2 \geq \frac{n^2}{k} \geq \frac{n^2}{d}.$$

In fact, the sum $\sum_{i=1}^{n} \lambda_i^2$ is equal to the trace of $A^T A$, which we can compute explicitly to be $\sum_{i,j} a_{i,j}^2$. Hence we have that

$$\sum_{i,j} a_{i,j}^2 \geq \frac{n^2}{d}.$$

But because $|a_{i,j}| \leq \sqrt{\epsilon}$ for $i \neq j$, we can bound the left hand side:

$$n + n(n-1)\epsilon \geq \left(\sum_{i=1}^{j} a_{i,i}^2\right) + \left(\sum_{i \neq j} a_{i,j}^2\right) = \sum_{i,j} a_{i,j}^2.$$

We obtain, therefore, that

$$n + n(n-1)\epsilon \geq \frac{n^2}{d},$$

and it clearly follows that

$$\epsilon \geq \frac{n-d}{d(n-1)}.$$

□

Using this lemma, we can now prove our lower bound on ϵ.

Theorem 3.3. *For any choice of n and N, the minimum value of ϵ satisfies*

$$\epsilon \geq \frac{N-n}{n(N-1)}.$$

Proof. Pick any collection of N unit vectors $v_1, \ldots, v_N$ in $\mathbb{R}^n$, and define ϵ as usual. Recall the definitions of A and G:

$$A = \begin{pmatrix} | & | & \cdots & | & | \\ v_1 & v_2 & \cdots & v_{N-1} & v_N \\ | & | & \cdots & | & | \end{pmatrix}$$

$$G(v_1, \ldots, v_N) = A^T A = \begin{pmatrix} v_1 \cdot v_1 & v_1 \cdot v_2 & \ldots & v_1 \cdot v_n \\ v_2 \cdot v_2 & v_2 \cdot v_2 & \ldots & v_2 \cdot v_n \\ \vdots & \vdots & \ddots & \vdots \\ v_n \cdot v_1 & v_n \cdot v_2 & \ldots & v_n \cdot v_n \end{pmatrix}$$

Let d be the rank of G. Since A has rank at most n, so does A^T; since the image space of G must be contained in the image space of A^T, we have that $d \leq n$. Note that G is real and symmetric, with 1s along the diagonal and all other entries with absolute value bounded by $\sqrt{\epsilon}$. We can apply the lemma to conclude

$$\epsilon \geq \frac{N-d}{d(N-1)}.$$

Since $d \leq n$, we have as desired

$$\epsilon \geq \frac{N-n}{n(N-1)}.$$

□

We present now a table of the bounds the theorem gives us for varying n and N:

	$n=2$	$n=3$	$n=4$
$N=2$	0	–	–
$N=3$	1/4	0	–
$N=4$	1/3	1/9	0
$N=5$	3/8	1/6	1/16
$N=6$	2/5	1/5	1/10
$N=8$	3/7	5/21	1/7
$N=12$	5/11	3/11	2/11

Table 1: Lower bounds on ϵ for different values of (n, N), given by Theorem 3.3.

Note that the bounds for $n = 2$ are not tight, as we proved in Section 2 that the minimum ϵ is given by $\cos^2(\pi/N)$ when $n = 2$. We shall see, however, that the bounds for ϵ are in fact tight in the cases where $N = n + 1$ and where $n = 3, N = 6$. One direction for further research might be checking whether the bounds provided by Theorem 3.3 are indeed attainable for other combinations of n and N.

Leon wrote this section. Deepak made minor edits to this section.

4 Some optimal configurations

Below we provide choices of n and N for which we can achieve the lower bound given by Theorem 3.3.

4.1 $n = 3, N = 6$

Let us consider collections of six vectors in $\mathbb{R}^3$. As can be seen in Table 1, the lower bound provided in this case by Theorem 3.3 is $\frac{1}{5}$ as well. We shall now see, in fact, that this lower bound is attainable.

Theorem 4.1. *For $n = 3$ and $N = 6$, the optimum value of ϵ is given by $\frac{1}{5}$.*

Proof. Consider the following 6 vectors on the unit sphere in $\mathbb{R}^3$ with variables a and b.

$$(a, b, 0), (a, 0, b), (b, 0, a)$$

$$(a, -b, 0), (a, 0, -b), (-b, 0, a)$$

There are three possible dot products we get from computing dot products between non-equal vectors chosen from these six vectors: $ab, -ab, a^2 - b^2$. This gives us two possible values of ϵ: $(a^2 - b^2)^2, (ab)^2$. It would seem that an optimal configuration can be obtained by setting $a^2 - b^2 = ab$. If we add in the constraint that $a^2 + b^2 = 1$ (since the six vectors we're interested in must live on the unit sphere), we get

$$a = \sqrt{\frac{1}{10}(5 - \sqrt{5}}$$

$$b = -\frac{1 + \sqrt{5}}{2}\sqrt{\frac{1}{10}(5 - \sqrt{5}}.$$

If we compute $\epsilon = (a^2 - b^2)^2 = (ab)^2$, we see that $\epsilon = \frac{1}{5}$. Coupled with our knowledge that $\frac{1}{5}$ is a lower bound on ϵ, we conclude that these six points give us an optimal configuration. □

4.2 $N = n + 1$

Let us now consider the situation in which we try to minimize ϵ for $n + 1$ vectors in an n dimensional space. From Theorem 3.3 we know that $\epsilon \geq \frac{1}{n^2}$, so it suffices to show there exist some $n + 1$ vectors in n-space that achieve an ϵ equal to $1/n^2$.

To do this, we first introduce the notion of an n-dimensional simplex.

Definition 4.2. *An n-dimensional simplex is given by the convex hull of the $n+1$ points in $\mathbb{R}^{n+1}$ described as having a 1 in a single coordinate and zeros in every other coordinate.*

Figure 5 shows diagrams of a 2-dimensional simplex. Notice that an n dimensional simplex is constructed in $n+1$ dimensions, but exists in a hyperplane of $\mathbb{R}^{n+1}$.

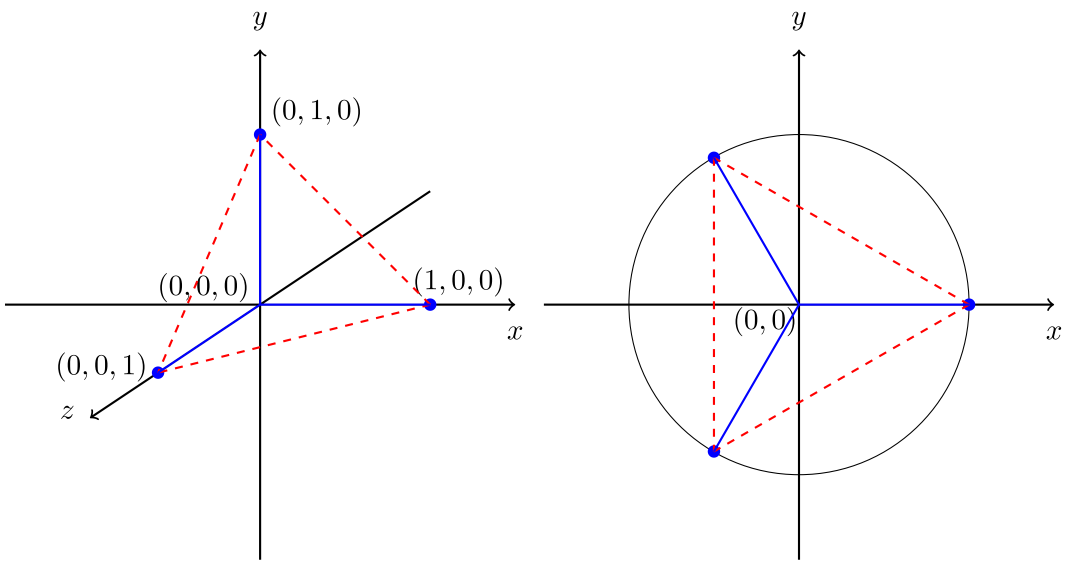

Figure 5: A 2-dimensional simplex drawn in $\mathbb{R}^3$ on the left, and then centered at the origin in $\mathbb{R}^2$ on the right.

Before proceeding with the theorem, we notice that when $n = 2$, the configuration of $N = 3$ vectors that produced the minimum value of ϵ made up a simplex centered at the origin (an equilateral triangle).

Theorem 4.3. *If $N = n + 1$, then the minimum value of ϵ for $\frac{1}{n^2}$. This ϵ is achieved when the vectors are arranged in an n dimensional simplex centered at the origin.*

Proof. Let ϵ be defined as before for a given set of unit vectors $v_1, v_2, \dots, v_{n+1}$. We show that the origin-centered simplex configuration of the vectors $v_1, \cdots v_{n+1}$ produces a value $\epsilon = \frac{1}{n^2}$.

We must first describe the vectors that make up the origin centered simplex. To center our simplex at the origin we take each vector in the original simplex, and subtract from it the vector representing the centroid of the original simplex. The centroid of the original simplex takes the form $(\frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N})$, so the points of the new simplex take the form

$$\left(-\frac{1}{N}, \cdots, -\frac{1}{N}, \frac{N-1}{N}, -\frac{1}{N}, \cdots, -\frac{1}{N}\right).$$

	$n=2$	$n=3$	$n=4$
$N=2$	0.0	–	–
$N=3$	0.25	0.0	–
$N=4$	0.5	0.11	0.0
$N=5$	0.66	0.21	0.07
$N=6$	0.75	0.22	0.14
$N=7$	0.82	0.36	0.18
$N=8$	0.86	0.44	0.24
$N=9$	0.89	0.49	0.26
$N=10$	0.91	0.51	0.32
$N=11$	0.93	0.55	0.35
$N=12$	0.94	0.64	0.41

Table 2: Minimum values of ϵ for different values of (n, N), as computed by our numerical method

---

If we renormalize these vectors so that each vector has unit length, we get

$$\sqrt{\frac{N}{N-1}}\left(-\frac{1}{N}, \dots, -\frac{1}{N}, \frac{N-1}{N}, -\frac{1}{N}, \dots, -\frac{1}{N}\right).$$

Now, taking the dot product between any two such vectors gives us

$$\frac{N}{N-1}\left(-2 \cdot \frac{N-1}{N^2} + (N-2) \cdot \frac{1}{N^2}\right) = \frac{N}{N-1} \cdot -\frac{1}{N} = -\frac{1}{N-1} = -\frac{1}{n}.$$

So for this configuration $\epsilon = \frac{1}{n^2}$. By Theorem 3.3 we know that $\epsilon \geq \frac{1}{n^2}$, so we have achieved the lower bound. □

Yajit wrote this section. Leon edited this section. Deepak made minor edits and produced the figures.

5 Numerical analysis and associated conjectures

Because of the difficulty involved with visualization in dimensions 3 and above, we used a computational approach to compute near-optimal epsilon values for different (n, N) pairs. We describe this method in greater detail below.

Our general intuition tells us that perturbing a set of nearly optimal vectors slightly could give us a new set of vectors that produce even a smaller epsilon value. Given this, we start off with a set of random vectors, and try to move these vectors towards smaller epsilon values.

At every iteration, we try adding a vector whose norm becomes smaller with every passing iteration, to the already computed optimal vector. We present pseudocode for this algorithm below.

Though this approach seems simple, it produces results which seem to agree with the ϵ we derived for $(2, N)$ in Section 2, and with the ϵ we derived for $(3, 6)$ and $(n, n+1)$ in Section 4. This gives us some confidence in the accuracy of our algorithm. However, it must be noted that the data does not agree with the bounds produced by Theorem 3.3. There are multiple interpretations for this fact. First, it is possible that the lower bounds produced by Theorem 3.3 are not always tight – this is easy to see for the case of $(2, N)$. Alternatively, it is possible that our code simply becomes inaccurate for larger n and N. This could be a fruitful direction for further research.

As another note, in some instances the algorithm sometimes finds extrema that are not the extrema we have identified in cases where we provide optimal configurations. For example, for $n = 3, N = 4$, the ϵ value attained is very close to the optimal epsilon value of $\frac{1}{9}$, however the corresponding configuration is not a tetrahedron. This implies that in some cases there may be multiple configurations of vectors that are not equivalent up to rotation yet produce the optimal ϵ value.

```python
def get_min_epsilon(n, N, num_iter, start_vectors):
    min_epsilon = 1.0  # Start off with the worst possible epsilon
    current_vectors = start_vectors
    i = 0
    while (i < num_iter):
        temperature = 1.0 / float(i + 1)

        # Get perturbing vectors
        random_vectors = get_random_vectors(n, N)

        # Now perturb current_vectors
        new_vectors = list()
        for j in xrange(N):
            new_vector = list()
            for k in xrange(n):
                new_vector.append(
                    current_vectors[j][k] + (
                        temperature * random_vectors[j][k]))
            new_vectors.append(new_vector)
        normalize_vectors(new_vectors)

        # Accept change only if new_vectors produces a better epsilon
        epsilon = compute_epsilon(new_vectors)
        if epsilon < min_epsilon:
            min_epsilon = epsilon
            current_vectors = new_vectors
            i += 1

    return min_epsilon, current_vectors
```

Algorithm: The numerical method used to generate optimum epsilon values – note that we call this method multiple times with different start_vectors to obtain a reasonable epsilon estimate

---

Deepak wrote this section. Leon edited this section. Yajit edited this section.

6 A variation

We now consider a variation of the above problem, where we actually try to place N unit vectors in n-space as far apart as possible from each other. More concretely, we define δ as

$$\delta = \max_{i \neq j} v_i \cdot v_j,$$

and we formalize our problem as finding the N vectors $v_1, v_2, \ldots, v_N$ that minimize δ.

Our results on this section concern the case when $N = n + 1$. We shall see, in fact, that the optimal value of δ is given by $-\frac{1}{n}$. To begin, we introduce the paralleogram law.

6.1 The Parallelogram Law

First we give the definition of the norm of a sum of vectors in $\mathbb{R}^n$ in terms of the dot product:

Definition 6.1. *For a vector $v \in \mathbb{R}^n$ the norm of v is the positive value of $||v||$ given by the equation*

$$||v||^2 = v \cdot v.$$

Theorem 6.2 (Parallelogram Law)**.** *For vectors $v_1, \cdots, v_n \in \mathbb{R}^n$, the following identity holds*

$$||v_1 + \cdots + v_n||^2 = \sum_{i=1}^{n} ||v_i||^2 + 2 \sum_{1 \leq i < j \leq n} v_i \cdot v_j.$$

Proof. First consider the following identity from the definition of the norm for vectors $u, v \in \mathbb{R}^n$:

$$||u + v||^2 = (u + v) \cdot (u + v) = u \cdot u + u \cdot v + v \cdot u + v \cdot v = ||u||^2 + ||v||^2 + 2u \cdot v.$$

Now if we let $u = v_1$ and $v = v_2 + \cdots + v_n$ we get

$$||v_1 + v_2 + \cdots + v_n||^2 = ||v_1||^2 + ||v_2 + \cdots + v_n||^2 + 2 \sum_{i=2}^{n} v_1 \cdot v_i.$$

If we recurse on $||v_2 + \cdots + v_n||^2$ using induction we get the desired result.

6.2 Variation: Minimizing the dot product

Given the above bound, we proceed to state and prove the following theorem.

Theorem 6.3. *If $N = n + 1$ and the vectors in question are denoted $v_1, \cdots, v_N$, then over all configurations of these vectors the minimum value of $v_i \cdot v_j$ for $i \neq j$ is $\frac{-1}{n}$. This is achieved when the vectors are arranged in an n dimensional simplex centered at the origin.*

Proof. Let $\delta = \max_{i \neq j}\{v_i \cdot v_j\}$ for a given set of unit vectors $v_1, v_2, \ldots, v_{n+1}$, and let δ_{min} be the minimum possible δ obtained across all unit vectors $v_1, v_2, \ldots, v_{n+1}$. We computed in the proof of Theorem 4.3 that for the origin centered simplex, we obtain a value of $\delta_{\min} = \frac{-1}{n}$, so $\delta_{\min} \leq \frac{-1}{n}$. So we need only show that $\delta_{min} \geq -\frac{1}{n}$ as well.

To see that $\delta_{min} \geq -\frac{1}{n}$ we use the parallelogram law.

Recall that the parallelogram law states that

$$||v_1 + \cdots + v_{n+1}||^2 = \sum_{i=1}^{n+1} ||v_i||^2 + 2 \sum_{1 \leq i < j \leq n+1} v_i \cdot v_j.$$

In our case we know three things. First, since the norm of any vector is greater than or equal to 0, we can conclude that $||v_1 + \cdots + v_{n+1}||^2 \geq 0$. Second, $v_i \cdot v_j \leq \delta$ for any $i < j$. Third, $||v_i|| = 1$. Therefore

$$0 \leq n + 1 + n(n+1)\delta.$$

This implies that $\delta \geq -\frac{1}{n}$ for any set of unit vectors $v_1, v_2, \ldots, v_{n+1}$, from which we can conclude that $\delta_{min} \geq -\frac{1}{n}$ as well.

Since $\delta_{min} \leq -\frac{1}{n}$ and $\delta_{min} \geq -\frac{1}{n}$, we can conclude that $\delta_{min} = -\frac{1}{n}$. □

Yajit wrote this section. Deepak made edits to this section.