MATH598B Homework 1

Alex Fruge

February 11, 2025

Suppose we have a vector space $\mathbb{R}^d$, with n vectors in this space.

1 Question 1

What is the maximum number n of orthogonal vectors we can put into this vector space? Explain why, in detail.

Answer: the maximum number of orthogonal vectors that can fit into $\mathbb{R}^d$ is d.

Why?:

For two vectors to be orthogonal, their inner product must be zero. In order for a group of vectors to be orthogonal, this relation must hold true for all pairs of vectors within the group. One useful property we get when assuming that the vectors are orthogonal is that they are also linearly independent, which helps us set the upper bound for the number of vectors in the vector space.

Since $\mathbb{R}^d$ is a d-dimensional space, the maximum number of linearly independent vectors in $\mathbb{R}^d$ is d. Since orthogonal vectors are also linearly independent, it also follows that the maximum number of orthogonal vectors in $\mathbb{R}^d$ is d.

An example of this is the standard basis vectors for $\mathbb{R}^d$.

$$e_1 = (1, 0, 0, \dots, 0), \quad e_2 = (0, 1, 0, \dots, 0), \quad \dots, \quad e_d = (0, 0, 0, \dots, 1)$$

These vectors satisfy $e_i \cdot e_j = 0$ for $i \neq j$, which means that there are d orthogonal vectors in $\mathbb{R}^d$.

2 Question 2

Suppose we concatenate the n orthogonal vectors, each in $\mathbb{R}^d$, into a single matrix. What properties will this matrix have? Be thorough.

Answer: Let the matrix A be defined by concatenating n orthogonal vectors in $\mathbb{R}^d$. If the vectors are normalized, then they form an **orthonormal** set ($v_i^\intercal v_j = 0$ for $i \neq j, v_i^\intercal v_i = 1$). If this is the case, then we can say that A is **semi-orthogonal**, which means that $A^\intercal A = I_n$ where I_n is the $n \times n$ identity matrix.

Since the columns are orthogonal and non-zero, they are **linearly independent**, which means that **rank** $A = \mathbf{min}(n, d)$ If $n \leq d$.

(*Note:* if $n = d$, and the columns are orthonormal, then A is orthogonal)

3 Question 3

Suppose $n \gg d$. Assuming they are packed "optimally", what bounds can we place on the expected dot product of two random vectors from our set? To answer this, you should do some mix of reading papers, trying to derive a bound, and doing numerical experiments. Include your code, cite your sources.

In this context, for the vectors to be packed optimally, I'll be assuming that they are uniformly distributed in the unit sphere in $\mathbb{R}^d$. In this situation, we can take two independently sampled vectors from our space, and find the expected value of their dot product.

I'll be using a simulation (`https://github.com/alexfruge/decoding-gpt-fruge/blob/main/Homework/hw1sim.py`) to gain some initial intuition as to what these bounds could be.

Dimension (d)	Mean Dot Product	Variance
10	$5.320\,471 \times 10^{-5}$	$2.040\,759 \times 10^{-7}$
50	$2.963\,093 \times 10^{-5}$	$3.253\,653 \times 10^{-8}$
100	$2.457\,548 \times 10^{-5}$	$1.105\,645 \times 10^{-8}$
500	$7.106\,165 \times 10^{-6}$	$5.118\,772 \times 10^{-9}$
1000	$7.432\,451 \times 10^{-6}$	$1.155\,743 \times 10^{-9}$

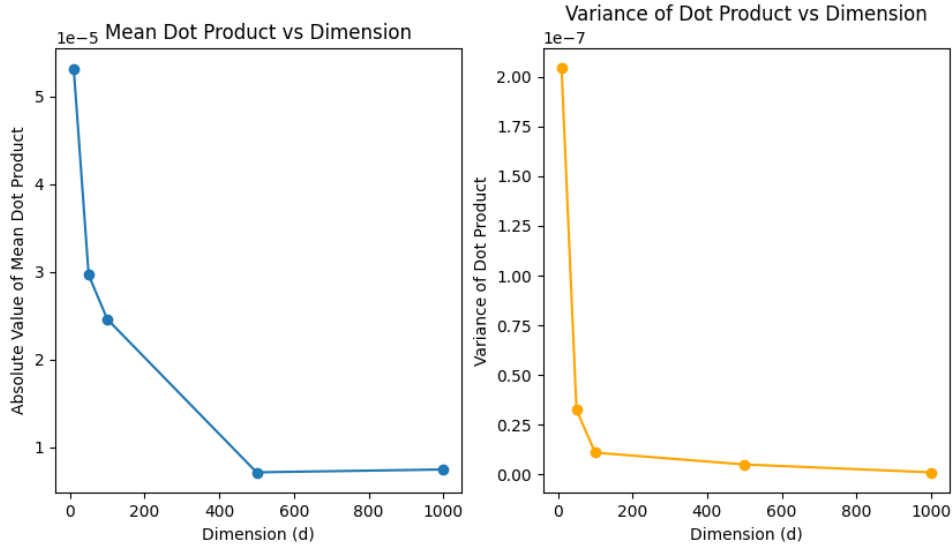Table 1: Mean Dot Product and Variance for Different Dimensions

Figure 1: Plot showing simulation results for mean dot product and variance of dot products as dimension changes.

4   Question 4

Why do you think thinking about this is relevant for the subject of the course? What do you think these vectors are supposed to represent, and why might we care about them being "optimally packed"?

Whenever the vectors in our space are "optimally packed", it means that

5 LLM Logs

Here's a prompt that I asked for question 3: Suppose we have a vector space $\mathbb{R}^d$, with n vectors in this space. Assume that $n \gg d$. What kind of bounds could be useful for finding the expected value of two random vectors from the set?

ChatGPT-4o Response: Given n vectors in $\mathbb{R}^d$ with $n \gg d$, a useful approach for finding the expected value of the dot product (or other function) of two randomly chosen vectors is to consider concentration inequalities and properties of high-dimensional geometry. Here are some useful tools and bounds:

Concentration of Measure: In high-dimensional spaces, random vectors tend to be nearly orthogonal. Specifically, if the vectors are normalized, the dot product between two randomly chosen vectors is close to zero with high probability. This can be formalized using concentration inequalities like Hoeffding's inequality or McDiarmid's inequality, depending on the dependence structure.

Johnson-Lindenstrauss Lemma: This lemma states that a set of n points in $\mathbb{R}^d$ can be embedded into $\mathbb{R}^k$ (with $k = O(\log n / \epsilon^2)$) such that pairwise distances are approximately preserved. This can be used to bound the expected distance or dot product between two vectors, as the reduced dimensionality preserves the geometric structure.

Inner Product Expectation: If the vectors are drawn uniformly from the unit sphere S^{d-1}, then the expected dot product between two randomly chosen vectors is:

$$\mathbb{E}[\langle X, Y \rangle] = 0$$

because of symmetry. The variance is $\frac{1}{d}$, suggesting that the dot products are typically small when d is large.

Random Matrix Theory: If the vectors are organized into a matrix A of size $d \times n$, then properties of the Gram matrix $G = A^{\intercal}A$(e.g., spectral norm bounds) can provide information about the correlations between the vectors.

Covering Numbers and Metric Entropy: For understanding the distribution of the vectors, bounds on covering numbers or metric entropy can provide insight into the expected geometric arrangement of the vectors, which can, in turn, help estimate expected values of distances or dot products.

If you can provide more specifics about the distribution of the vectors or the exact expected value you're looking for, I can tailor the bounds and tools more precisely!

Prompt intent: I intended for this prompt to be a little more general, but 4o went into a good bit of detail, which ended up being useful to kickstart my own research into the problem.