Student Name: Minh Phan
Course: AMATH 352
Email address: minhphan@uw.edu

Homework 3

---

1 Introduction and overview of the problem

In linear algebra, the concept of linear independence is vital because a set of linear independent vectors define spaces and spans. To put it differently, by identifying independent vectors, we can describe the other dependent vectors based on them and henceforth retaining the vital properties that define a space. Applying the concept to practical fields like data science or machine learning, we can efficiently drop the dependent data rows without loss of much information contained in the data. [1]

In this report, I want to investigate the linear dependence of certain collections of vectors, include both randomly generated collections (reproducibility added using seeds) and non-randomly generated collections. Another problem that we also want to tackle in this report is the effectiveness of the Gram-Schmidt process to orthogonalize sets of vectors. A more detailed description of the process is provided in the next section. Testing the effectiveness of a method or process like this is important as it helps readers to gain insight into the processes themselves and propels enhancement/developments of newer, more effective ones.

2 Theoretical background and description of algorithm

The Gram-Schmidt process is a procedure which takes a non-orthogonal set of linearly independent vectors and constructs an orthonormal basis that span the same space spanned by the original set itself [4]. To understand better the Gram-Schmidt process, I provide some theoretical background of some of the concepts it relates to:

- Linear independence

  Mathematically, a set of vectors $v_1, v_2, ..., v_k$ is defined as linearly independent if the equation

  $$x_1\underline{v}_1 + x_2\underline{v}_2 + ... + x_k\underline{v}_k = 0$$

  has only one trivial solution $x_1 = x_2 = ... = x_k = 0$ and linearly dependent otherwise. [3]

- 2-norm

  The 2-norm represents the length of a vector. For a vector $\underline{x}$, the 2-norm vector can be represented as $||\underline{x}||_2$ or shortened as $||\underline{x}||$. The method to calculate the 2-norm vector is:

  $$||\underline{x}|| = ||\underline{x}||_2 = \sqrt{x_1^2 + x_2^2 + ... + x_k^2}$$

- Orthogonality

  Two vectors $\underline{x}, \underline{y}$ are said to be orthogonal or perpendicular if the dot product of them is 0. [3] According to the theorem in lecture 8, we can also prove that orthonormal vectors are linearly independent.

- Dot product

  The dot product measures angles between vectors and computes the length of a vector. The dot product of two vectors is a scalar. Within the context of the process, one noticeable property we have is that the dot product of a vector with itself is length of the vector squared. $\underline{x} \cdot \underline{x} = ||\underline{x}||^2$

- Projection and projection operator

  $$proj_u(v) = \frac{\underline{u} \cdot \underline{v}}{\underline{u} \cdot \underline{u}}\underline{u}$$

  The vector projection of a vector $\underline{v}$ on a nonzero vector $\underline{u}$ is the orthogonal projection of the former onto a straight line parallel to the latter.

In Margalit and Rabinoff's version, the resulted set after the Gram-Schmidt process is not normalized, but we can perform normalization after each main step to better fit our computation. [3] In short, let us have a vector set $\underline{u}_1, \underline{u}_2, ..., \underline{u}_k$, it has three main steps:

1. Step 1: Normalize the first vector in the set

$$\underline{v}_1 = \underline{u}_1, \underline{e}_1 = \frac{\underline{v}_1}{||\underline{v}_1||}$$

2. Step 2: Subtract the projection of the second vector on the first vector from the original, then normalize it

$$\underline{v}_2 = \underline{u}_2 - proj_{\underline{v}_1}(\underline{u}_2), \underline{e}_2 = \frac{\underline{v}_2}{||\underline{v}_2||}$$

3. Step 3: Subtract the projection of the third vector on the first vector v_1 and the its projection on the second orthogonal vector v_2 from the original third vector in the set, then normalize it just like in step 2. Repeat the same procedure for the rest of the vectors.

$$\underline{v}_3 = \underline{u}_3 - proj_{\underline{v}_1}(\underline{u}_3) - proj_{\underline{v}_2}(\underline{u}_3), \underline{e}_3 = \frac{\underline{v}_3}{||\underline{v}_3||}$$

$$...$$

$$\underline{v}_k = \underline{u}_k - \sum_{j=1}^{k-1} proj_{\underline{v}_j}(\underline{u}_k), \underline{e}_k = \frac{\underline{v}_k}{||\underline{v}_k||}$$

3 Computational Results

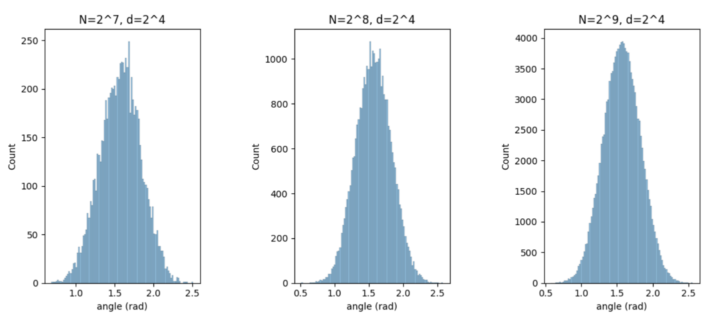

Figure 1: Pairwise angle distribution in collections of N $\mathbf{R}^d$ vectors, $a_i \in N(0, 1)$

In Figure 1 we can see that the distribution of all collections are bell-curved (normally distributed) and centered around value of $\pi/2$, which means that many of the vectors in these collections are perpendicular (orthogonal) to each other. Based on this fact we can say that their linear independence are strong (but not completely linearly independent).

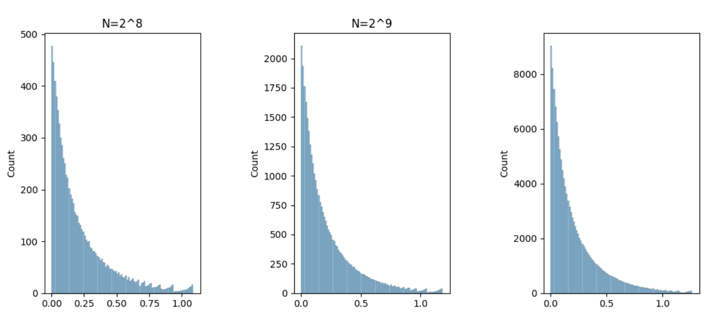

Figure 2: Pairwise angle distribution collection of N $\mathbf{R}^d$ vectors, $b_{ij} = \frac{1}{i+j-1}$

Unlike the previous figure, the distribution of non-random collections of vectors' angles of pairwise vectors shown in Figure 2 below are heavily right-skewed and have all modes at around 0 rad. This means that not many of the vectors are orthogonal to each other, and from the fact from lecture 8 I have mentioned above, the linear independence of these sets are substantially weak. These vectors are heavily dependent on each other.

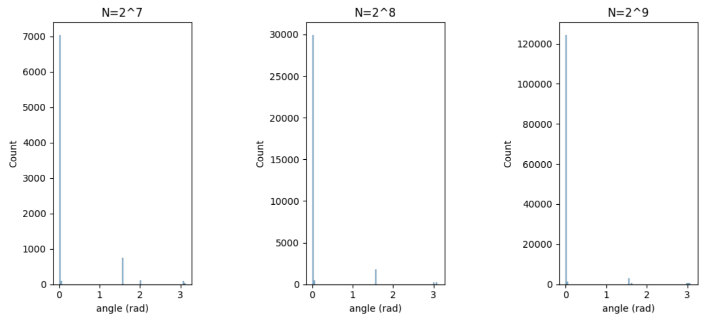

Figure 3: Pairwise angle distribution of orthogonalized sets of N $\mathbf{R}^d$ vectors, $b_{ij} = \frac{1}{i+j-1}$

From figure 3 we can see that all of them have a handful of angles at around $\pi/2$, but considerable numbers of angles close to 0. When we apply it on linearly dependent sets, the Gram-Schmidt process can be used to identify the linear dependence of the sets by comparing the number of orthogonalized vectors whose pairwise angles return $\pi/2$ to the number of vectors in the original set.

However, the facts that there are so many close-to-zero angles after performing the process shows that it does not work as expected. Indisputably we cannot find the angle between zero vectors and therefore if there are so many close-to-zero angles like this it probably means that we have serious rounding errors for a big part of the q_i's. I can also eliminate the reasoning that there is any issues in my implementation as we can still see a handful of $\pi/2$ angles in all of the graphs and not spread out to any other values, and the only outlier values is the range very close to 0. This result is expected. According to an article on ScienceDirect, the results yields poor results when some of the vectors are almost linear dependent, which is what we see here in Figure 2 and 3. The rounding errors happen because in the classical model, we subtract the projections all at once. [2]

4 Summary and Conclusions

In summary, we can see that sets of random Gaussian vectors normally have strong linear independence, and that non-random sets of vector that rely on orders of the vectors and positions of elements in the vectors themselves are very likely to be linearly dependent.

The Gram-Schmidt process is a notable one in linear algebra to orthogonalize vector sets and build up other important concepts in the field like bases, spans, dimensions, etc. Applying and evaluating the classical Gram-Schmidt process, we can conclude that it is numerically unstable, primarily due to apparent rounding errors. Therefore, using the modified Gram-Schmidt process, by removing the projections once at a time, is a better option to solve the issue. [2]

References

[1] Mike Beneschan. *How To Understand Linear Independence (Linear Algebra)*. Medium, July 2020. URL: `https://medium.com/swlh/how-to-understand-linear-independence-linear-algebra-8bab1d918509` (visited on 10/22/2022).

[2] William Ford. *Gram-Schmidt Orthonormalization*. www.sciencedirect.com, 2015. URL: `https://www.sciencedirect.com/topics/mathematics/gram-schmidt-process` (visited on 10/24/2022).

[3] Dan Margalit and Joseph Rabinoff. *Interactive Linear Algebra*. 2019. URL: `https://textbooks.math.gatech.edu/ila/ila.pdf`.

[4] Eric W. Weisstein. *Gram-Schmidt Orthonormalization*. MathWorld–A Wolfram Web Resource. URL: `https://mathworld.wolfram.com/Gram-SchmidtOrthonormalization.html` (visited on 10/24/2022).