

Linear Algebra HW 4

gjd9961

September 2021

Name: Giulio Duregon

1 Problem 4.1

a) $N(\cdot)$ is a norm on \mathbb{R}^2 if it satisfies the following conditions:

- Homogeneity $\|\alpha v\| = |\alpha| \times \|v\|$ for all $\alpha \in \mathbb{R}$ and $v \in \mathbf{V}$
- Positive Definiteness: if $\|v\| = 0$ for some $v \in \mathbb{R}$, then $v = 0$
- Triangular inequality: $\|u + v\| \leq \|u\| + \|v\|$ for all $u, v \in \mathbf{V}$

Lets see if our norm satisfies these conditions, starting with homogeneity.

Let (x_1, x_2) be the entries in the 2-dimensional vector x with $|x_1| > |x_2|$

$$\alpha \times x = (\alpha x_1, \alpha x_2) \rightarrow N(\alpha x) = \max(|\alpha x_1|, |\alpha x_2|) = |\alpha| \times |x_1| = |\alpha| \times N(x) \text{ Satisfies Homogeneity}$$

We can check positive definiteness by testing different vectors on our norm. If any value in the vector we apply to the norm to are non-zero, the output of the $N(\cdot)$ will be a positive number. If the entries of x are 0, that is the only case where $N(x) = 0$. Let $z, x, y \in \mathbb{R}^2$ and $z = (0, 0), x = (-5, 10), y = (-5, -1)$

$$N(z) = N(z) = \max(|0|, |0|) = 0 \text{ Satisfies Positive Definiteness}$$

$$N(x) = N(x) = \max(|-5|, |10|) = 10 \text{ Satisfies Positive Definiteness}$$

$$N(y) = N(y) = \max(|-5|, |-1|) = 5 \text{ Satisfies Positive Definiteness}$$

Lastly, we need to check the triangle inequality, that $\|u + v\| \leq \|u\| + \|v\|$. Let $u = (-10, 2)$ and $v = (2, 4)$ therefore $u + v = (-8, 6)$

$$\|u + v\| = \max(|-8|, |6|) = 8 \leq 14 = 10 + 4 = \max(|-10|, |2|) + \max(|2|, |4|) = \|u\| + \|v\|$$

Therefore $\|u + v\| \leq \|u\| + \|v\|$ the triangular inequality holds and we have a valid norm.

b) $x = (8, 0)$ and $y = (7, 7)$ Using $N(\cdot)$ which of x and y has the bigger norm? Using the Euclidean norm which of x and y has the bigger norm?

- $N(\cdot)$ Norm: $N(x) = \max(|8|, |0|) = 8$ and $N(y) = \max(|7|, |7|) = 7$. Norm $N(x) > N(y)$
- Euclidian Norm: $\|x\|_2 = \sqrt{8^2 + 0^2} = \sqrt{64} = 8$ while $\|y\|_2 = \sqrt{7^2 + 7^2} = \sqrt{98} \approx 9.9$
 $\|y\|_2 > \|x\|_2$, Euclidean norm of y is bigger than Euclidean norm of x .

c) To represent and rank students grades in the class, I would rather use the Euclidean norm than the $N(\cdot)$ norm we have defined. The reason is, if we used the $N(\cdot)$ norm, we would just be taking the maximum of each students scores for the mid term and the final, and if one student bombed the midterm and did well on the final, he would outrank one that did well on both, but not great on either. The Euclidean norm on the other hand would essentially give us the score vectors "length" away from the origin, with higher scores pushing the vector farther and farther away from the origin, and thus increasing its norm. The Euclidean norm yields length, which useful to measure the spatial difference in students scores and thus a good way to rank them from one another.

2 Problem 4.2

Any inner product must have the property that $\langle x, y \rangle = \langle y, x \rangle$. Therefore, our inner product defined by $\langle \cdot, \cdot \rangle_A$ by

$$\langle x, y \rangle_A = x^\top A y = \sum_{i,j=1}^N x_i A_{ij} y_j$$

where $A \in \mathbb{R}^{n \times n}$ must have this property as well that $\langle x, y \rangle = \langle y, x \rangle$. Lets show how a,b,c and must follow for our inner product to be valid.

a) The matrix A is symmetric. This follows from the property that for an inner product to be valid, it must satisfy symmetry, and that $\langle x, y \rangle = \langle y, x \rangle$. Lets ensure that this holds:

$$\langle x, y \rangle = x^\top A y = y^\top A x \rightarrow (x^\top A y)^\top = y^\top A \rightarrow y^\top A^\top x = y^\top A x$$

the only way for this equality to hold is if $A = A^\top$ which by definition means that A is symmetrical. Therefore, for our inner product to be valid, A must be symmetrical.

b) I'm going to prove that the matrix A must have positive diagonal values, $A_{ii} > 0$ for $i = 1, \dots, n$ by counter example. Assume our (symmetric) matrix A doesn't need to have positive entries in the diagonal while still satisfying all of the properties of an inner product.

$$\text{Let } \vec{x} \in \mathbb{R}^2 \text{ where } x = (1, 1) \text{ and let } A \in \mathbb{R}^{2 \times 2} \text{ where } A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\text{Therefore: } x^\top A x = x^\top \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = x^\top \begin{pmatrix} -1 \\ 1 \end{pmatrix} = (1 \quad 1) \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1 * 1 + 1 * 1 = 0$$

The above computation violates the property of Positive Definiteness, that when taking the inner product of a vector with itself, the result must be greater to 0 (and only 0 if the vector itself is the zero vector). Positive Definiteness: $\langle x, x \rangle \geq 0$ with equality only if $x = \emptyset$. We have found that if our matrix has negative entries in the diagonal, there are cases where Positive Definiteness is violated, and we don't have an inner product. We can also show that if there are diagonal entries equal to zero in our matrix we can have a non-zero x vector that violates positive definiteness. Again, the matrix A must have $A_{ii} > 0$ for all $i = 1, \dots, n$.

$$\text{Let } \vec{x} \in \mathbb{R}^3 \text{ where } x = (0, 1, 0) \text{ and let } A \in \mathbb{R}^{3 \times 3} \text{ where } A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\text{Therefore: } x^\top A x = x^\top \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = x^\top 0 = 0$$

c) For our inner product to be valid, A must be invertible. This statement follows from the combination of the problem statement (our matrix is square) and from the property of Positive Definiteness from inner products. To prove this, let's assume that A doesn't have to be invertible for our inner product to be valid.

We have our matrix $A \in \mathbb{R}^{n \times n}$ and we're going to assume that it has a kernel, $Ker(A)$, which follows because if a matrix is not invertible, it can't have full rank, and rank nullity tells us that if $Rank(A) < n$ then $dim(Ker(A)) \geq 1$. Then, the vector $\vec{x} \in Ker(A)$ where \vec{x} is not the zero vector, and that $A\vec{x} = \vec{0}$ where $\vec{0}$ is the zero vector with $1 \times n$ dimension with its entries all equal to 0. Then:

$$\langle x, x \rangle = x^T A \vec{x} \rightarrow x^T \times \vec{0} = 0$$

Positive Definiteness says that $\langle x, x \rangle = 0$ if and only if $x = 0$, however, we have just shown that $\langle x, x \rangle = 0$ with $x \neq 0$. Therefore, the property of Positive Definiteness is violated and we don't have an inner product.

In conclusion, if we want our inner product to be valid, our matrix A can't have a non-trivial Kernel. If there is no kernel, $Dim(Ker(A)) = 0$ and $Rank(A) = n$. We know our matrix is square ($A \in \mathbb{R}^{n \times n}$) and full-rank, therefore it is invertible.

3 Problem 4.3

a) Let a vector $y \in S$ and a vector $x \in V$ with its projection, $P_S(x) \in S$ with $\{v_1, \dots, v_n\}$ be the orthonormal basis for S . Therefore: y is a linear combination of the basis vectors of S : $y = \alpha_1 \times v_1 + \dots + \alpha_n \times v_n$ where $\alpha \in \mathbb{R}$, and also $P_S(x)$, is a product of the matrix transformation $SS^T x$ and can be expressed by a linear combination of basis vectors in S as follows: $P_S(x) = v_1 \langle v_1, x \rangle + \dots + v_n \langle v_n, x \rangle$. Lastly, we're going to be using the Euclidean dot product as our inner product $\langle x, y \rangle$

Firstly, let us calculate the dot product between the vectors x and y :

$$\text{Dot product of } x \text{ and } y: \langle x, y \rangle = \langle x, \alpha_1 \times v_1 + \dots + \alpha_n \rangle = \sum_{i=1}^n \alpha_i \langle x, v_i \rangle \quad (1)$$

Now let's calculate the dot product of $P_S(x)$ and y :

$$\langle P_S(x), y \rangle = \langle v_1 \langle v_1, x \rangle + \dots + v_n \langle v_n, x \rangle, \alpha_1 \times v_1 + \dots + \alpha_n \times v_n \rangle \quad (2)$$

Let $\omega = v_i \langle v_i, x \rangle$ and plug into our expression. We can now express the dot product as follows:

$$\sum_{i=1}^n \sum_{j=1}^n \omega_i \alpha_j \langle v_i, v_j \rangle \text{ where } \langle v_i, v_j \rangle = 0 \text{ if } i \neq j \text{ and } \langle v_i, v_j \rangle = 1 \text{ if } i = j \quad (3)$$

Since all the terms where $i \neq j$ are equal to 0, we can redefine our expression as:

$$\sum_{i=1}^n \omega_i \alpha_i = \sum_{i=1}^n \langle v_i, x \rangle \alpha_i \quad (4)$$

We can now see that y dotted with x is equal to y dotted with $P_S(x)$:

$$\langle y, P_S(x) \rangle = \sum_{i=1}^n \langle v_i, x \rangle \alpha_i = \sum_{i=1}^n \alpha_i \langle x, v_i \rangle = \langle x, y \rangle \quad (5)$$

And the equality holds!

b) $x - P_S(x)$ is orthogonal to S , as $P_S(x)$ and x and $x - P_S(x)$ form a triangle where $x - P_S(x)$ is the leg of the right triangle that is orthogonal and forms the 90 degree angle with the other leg. We can prove this rigorously using what we showed in part a).

Let $y, P_S(x) \in S$ and $x \in V$. If $x - P_S(x) \perp S$ then $\langle (x - P_S(x)), y \rangle = 0$

$$\langle (x - P_S(x)), y \rangle = 0 \text{ Begin with our equation} \quad (6)$$

$$\langle x, y \rangle - \langle P_S(x), y \rangle = 0 \text{ By property of inner products} \quad (7)$$

$$\langle x, y \rangle - \langle x, y \rangle = 0 \text{ Substitute } \langle x, y \rangle \text{ for } \langle P_S(x), y \rangle \text{ as we found in part a)} \quad (8)$$

$$0 = 0 \text{ End of proof} \quad (9)$$

We have shown that $\langle (x - P_S(x)), y \rangle = 0$ holds and that $x - P_S(x) \perp S$

c) Let $x \in V$ and $P_S(x) \in S$ be the vector x projected into the subspace S with $x - P_S(x) \in V$ the vector perpendicular to the subspace S that forms a triangle with x and $P_S(x)$.
Therefore: $\|x - P_S(x)\| \perp P_S(x)$

$$\|x - P_S(x) + P_S(x)\|^2 = \|x - P_S(x)\|^2 + \|P_S(x)\|^2 \text{ Pythagorean Theorem} \quad (10)$$

$$\|x\|^2 = \|x - P_S(x)\|^2 + \|P_S(x)\|^2 \text{ Simplify left side} \quad (11)$$

$$\|x\|^2 = \|x\|^2 - 2\langle x, P_S(x) \rangle + \|P_S(x)\|^2 + \|P_S(x)\|^2 \text{ expand the term } \|x - P_S(x)\|^2 \quad (12)$$

$$2\langle x, P_S(x) \rangle = 2 \times \|P_S(x)\|^2 \text{ subtract } \|x\|^2 \text{ from each side and simplify} \quad (13)$$

$$\|x\| \times \|P_S(x)\| \cos(\theta) = \|P_S(x)\|^2 \text{ subtract } \|x\|^2 \text{ divide by 2 from each side, expand dot product} \quad (14)$$

$$\|x\| \times \cos(\theta) = \|P_S(x)\| \text{ divide both sides by } \|P_S(x)\| \text{ and we're done} \quad (15)$$

We have found that the relationship between a vector and its projection can be described by the following:

$$\cos(\theta) \times \|x\| = \|P_S(x)\|$$

where $0 \leq \cos(\theta) \leq 1$ This indicates that at most $\cos(\theta) = 1$ then $P_S(x) = x$. Otherwise, the length of $\|P_S(X)\| < \|x\|$.

This makes sense as $\cos(\theta)$ can represent measure to how similar the vectors are, with a value of 1 meaning the vectors are on the same span and share the same direction, a value of 0 if they are orthogonal. If $\cos(\theta) = 1$ then the vectors projection is just the vector. If the vectors do not have $\cos(\theta)$ value of 1, then they are not on the same span, and $P_S(x) < x$.

4 Problem 4.4

Show that for any vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$: $(\sum_{k=1}^n x_k)^2 \leq n \sum_{k=1}^n x_k^2$.

Proof:

Let $x, y \in \mathbb{R}^n$ where $y_1, \dots, y_n = 1$ and $x_1, \dots, x_n \in \mathbb{R}$ and let $\|\cdot\|_2$ be the Euclidian Norm

$$|\langle x, y \rangle| \leq \|x\|_2 \times \|y\|_2 \text{ By Cauchy-Schwartz} \quad (16)$$

$$|\sum_{k=1}^n x_k \times 1| \leq \sqrt{\sum_{k=1}^n x_k^2} \times \sqrt{\sum_{k=1}^n 1} \text{ Expand Euclidian Norms} \quad (17)$$

$$|\sum_{k=1}^n x_k \times 1| \leq \sqrt{\sum_{k=1}^n x_k^2} \times \sqrt{n} \text{ Simplify far right summation } (1 \times n = n) \quad (18)$$

$$(\sum_{k=1}^n x_k)^2 \leq n \times \sum_{k=1}^n x_k^2 \text{ Square both sides} \quad (19)$$

We have arrived at our inequality and proved that it holds as long as Cauchy-Schwartz holds.