ASSIGNMENT 1: KEY

Exercise 1. (The Unit Ball) Recall the definitions of the sup norm $\|\cdot\|_{max}$, the Euclidean norm $\|\cdot\|_2$, the additive norm $\|\cdot\|_1$ and the p-norms $\|\cdot\|_p$ on $\mathbb{R}^n$. Draw diagrams of the closed unit ball in $\mathbb{R}^2$, $B_1 = \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$, under each of these norms. (Draw a separate diagram for $p > 1$ and $p < 1$ cases, note that for $p < 1$ $\|\cdot\|_p$ generates only a semi-norm (what fails?).)

Solution:

For the uniform norm, the diagram should show a square with vertexes at $(\pm 1, \pm 1)$.

For the 2-norm, the diagram should show a circle of radius 1 centered at the origin.

For the 1-norm, it should show a square with vertexes at $(\pm 1, 0)$ and $(0, \pm 1)$.

For the p-norm, $p > 1$, it should show a convex oval (thicker then a circle) such that as $p \to \infty$ the "circle" is approaching the square.

For $p < 1$, observe that it is not a norm (but a semi-norm), the triangle inequality does not hold, however, it is still possible to draw a diagram for the unit ball. It is not convex anymore.

Note: we did this exercise on the board in class, so if you still have problems, contact the instructor or any of the TAs for clarifications.

Exercise 2. (Euler's formula) Let $x \in \mathbb{R}$. Show that $e^{i\,x} = \cos x + i \sin x$.

Hint: Use the definition of the complex exponential from the Lecture Notes, i.e., the Taylor expansion at 0. The proof requires a rearranging of terms, but since the series is absolutely convergent we can make such a rearrangement.

Solution: We want to show that Euler's formula agrees with the power series of e^{ix}. Proceeding formally

$$\begin{aligned}\exp(ix) &= \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} \\ &= 1 + \frac{ix}{1!} + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \frac{(ix)^4}{4!} + ... \\ &= 1 + i\frac{x}{1!} - \frac{(x)^2}{2!} - i\frac{(x)^3}{3!} + \frac{(x^4)}{4!} + ...,\end{aligned}$$

where the terms alternate between real and exponential, and the coefficients alternate between positive and negative every two terms. Since the series is absolutely convergent, we may rearrange terms and factor out an i to get

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!}..... + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}...\right).$$

In closed form, this is

$$\sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{(2n)!} + i\sum_{m=0}^{\infty}(-1)^m \frac{x^{2m+1}}{(2m+1)!},$$

which is a power series expansion of $\cos x + i \sin x$.

Technically, it would be more correct to work only with the first $2n$ terms and show that the $2n$-th partial sum of the series $\exp ix$ can be broken into the first n terms of the sum defining $\cos x + i \sin x$. But this is what we are doing implicitly when we invoke absolute convergence to allow us to rearrange terms.

$$\sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{(2n)!} + i\sum_{m=0}^{\infty}(-1)^m \frac{x^{2m+1}}{(2m+1)!},$$

which is a power series expansion of $\cos x + i \sin x$.

Exercise 3. (Square) Show that for $z \in \mathbb{C}$, $z\bar{z} = |z|^2$. Do this using both the rectangular and polar representations of z (it should only take a couple of lines both ways). Is z^2 different from $|z|^2$, or $|z^2|$?

Solution:

- Using rectangular coordinates: let $z = x + iy$. Then $z\bar{z} = (x+iy)(x-iy) = x^2 - (iy)^2 = x^2 + y^2 = \mathrm{Re}(z)^2 + \mathrm{Im}(z)^2 = |z|$.

- Using polar (complex exponential) form: let $z = re^{i\theta}$. Then it is easy to check from Euler's formula and the even/odd properties of sine and cosine, that $\bar{z} = re^{-i\theta}$. Thus, we get $z\bar{z} = re^{i\theta}re^{-i\theta} = r^2e^{i(\theta-\theta)} = r^2 = |z|^2$.
- The easiest is to compare them in polar form: $z^2 = (re^{i\theta})^2 = r^2e^{2i\theta}$, $|z|^2 = |re^{i\theta}|^2 = r^2$, $|z^2| = |r^2e^{2i\theta}| = r^2$. So the first and last are the same and different from the second expression.

Exercise 4. (Linear Dependence and Bases) Prove statements, following the given steps:

(a) If $n+1$ vectors u_j (i.e., $u_1, u_2, ..., u_{n+1}$) lie in the span of n vectors v_l (i.e., $v_1, v_2, ..., v_n$), then the $n+1$ vectors u_j's are linearly dependent.

Sketch: Proceed by induction:

- *First, write out the problem for $n = 1$ and see that it is true. (You might also want to check $n = 2$ to understand the set up better.)*
- *For the inductive step, assume the conclusion when $n = k - 1$ (write it out).*
- *Now we will take $n = k$ and the goal will be to show that $u_1, u_2, ..., u_{k+1}$ are linearly dependent. Write each u_j in terms of the v_l's ($l = 1, ..., k$). Assume that the coefficient of v_k in the expansion of u_{k+1} is nonzero (can you justify this assumption? remember how Steven discussed on the board different cases why the coefficient in front of v_k can be chosen nonzero?). Using this, subtract a multiple of u_{k+1} from u_j to create a new set of k vectors w_j such that each w_j lies in the span of $\{v_1, ..., v_{k-1}\}$—note that in order for this to occur, the coefficients of v_k in the expansions need to cancel out.*
- *Using the inductive hypothesis, conclude that the w_j's are linearly dependent (there are only k of them!). Finally, show that this implies that the u_j's are linearly dependent (this again requires expanding in terms of the v_l's).*

(b) Let W be a subspace of a vector space V over $\mathbb{F}$. Let B_1 and B_2 be bases for W, and suppose that B_1 contains n elements. Then B_2 also contains n elements.

(*Hint: Use part (a).*)

(c) Let V be an n-dimensional vector space, and $v_1, ..., v_n$ be distinct vectors in V. Then $v_1, ..., v_n$ are linearly independent if and only if they span V.

(*Hint: Use part (a) for both implications.*)

Solution:

A. 1.

- *Base Step:* If $k = 1$, then u_1 and u_2 are both multiples of v_1, and hence are obviously linearly dependent.
- *Induction Step:* Assume that the result is true when $n = k - 1$. Let $v_1, ..., v_k \in V$ and $u_1, ..., u_k \in \text{span}\{v_1, ..., v_k\}$. By the definition of the span of a set of vectors, for each $j = 1, ..., k + 1$, there are scalars $a_{j,l}$ such that

$$u_j = a_{j,1}v_1 + ... + a_{j,k}v_k$$

If w_{k+1} is zero, then the u_j's are trivially dependent, so assume otherwise. Then $a_{k+1,l}$ is nonzero for some l, without loss of generality $l = k + 1$. For $j = 1, ..., k$ define

$$w_j = u_j - \frac{a_{j,k}}{a_{k+1,k}}w_{k+1}.$$

By definition, the coefficients of v_k in the expansion of w_j cancel out, so that each w_j lies in the span of $\{v_1, ...v_k\}$. By the inductive hypothesis, the w_j's are linearly dependent; that is, there are constants $c_1, ..., c_k$, not all zero, such that

$$\begin{aligned} 0 &= c_1w_1 + ... + c_kw_k \\ &= [c_1u_1 + ... + c_ku_k] + \frac{1}{a_{k+1,k}}[c_1a_{1,k} + ... + c_ka_{k,k}]u_{k+1}. \end{aligned}$$

There are two cases. Either the coefficient in front of u_{k+1} is nonzero, in which case the u_j's are by definition linearly dependent, or the coefficient is zero, in which case the term in brackets at the left of the second line is zero, and again the u_j's are linearly dependent because at least one c_j is nonzero. This concludes the induction step, and the lemma is proved.

B. We proceed by contrapositive. Let $\{w_j\}$ and $\{v_l\}$ be two finite sets spanning V, and suppose that they do not have the same number of elements. We show that it is not possible for both to be bases for V. Assume, without loss of generality, that $\{v_l\}$ is the smaller set, containing L elements. Since the v_l's span V, each w_j, lies in the span of the v_l's for $j = 1, ..., L+1$. By (A), this implies that the w_j's are linearly dependent, and hence cannot be a basis for V.

C. Let $u_1, ..., u_n$ be a basis for V. First suppose that $v_1, ..v_n$ are linearly independent. Let $w \in V$. Then by A and the fact that the u_j's span V, the collection $\{w, v_1, ..., v_n\}$ must be linearly dependent, so that there are constants c_k, not all zero, such that

$$c_0 w + c_1 v_1 + ... + c_n v_n = 0.$$

Since the v_k's are linearly dependent, c_0 must be nonzero, and hence, by dividing each c_k by c_0 and rearranging the equation, we can solve for w in terms of the v_k's, which shows that the v_k's span V.

Now suppose the v_k's span V. If they were linearly dependent, then some v_k would lie in the span of the others, and hence a collection of $n-1$ of the v_k's would span V. By (A), this would imply that the u_j's were linearly dependent, contradicting the definition of a basis.

Exercise 5. (Inner Product Space $\longrightarrow$ Normed Space $\longrightarrow$ Metric Space) Here we work in the opposite direction as the lecture notes on metrics, norms and inner products. Using the suggestions that follow, prove that **an inner product space is a metric space** by first proving the Cauchy-Schwartz inequality, and then showing that an inner product induces a norm and a norm induces a metric:

(a) (Cauchy-Schwartz Inequality) Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and let $\|\cdot\| : V \to \mathbb{R}$ be a norm induced by the inner product (recall what this means). Show that for any $x, y \in V$,

$$|\langle x, y \rangle| \leq \|x\| \, \|y\| .$$

Follow the steps below.

(i) First, suppose $y = 0$. Show that the inequality holds. For the remainder of the proof, we will assume $y \neq 0$.

(ii) For x, $y \in V$, define $F_{x,y} : \mathbb{C} \to \mathbb{C}$ by $F_{x,y}(t) = \langle x + ty, x + ty \rangle$. Expand F using linearity and complex linearity of the inner product. What inequality do we know $F_{x,y}$ satisfies?

(iii) Now—just for this part—assume we are dealing with a real inner product and $F_{x,y} : \mathbb{R} \to \mathbb{R}$. Use elementary calculus to find the only critical point of $F_{x,y}$,, that is, the point at which $F_{x,y}$ takes its minimum. [The solution is the minimum even when the inner product is complex, but it is not obvious that we can use methods of elementary calculus with a function of a complex variable.] Explain why the nonzero assumption on y is necessary.

(iv) Next we need the following: if $z \in \mathbb{C}$, then $z\bar{z} = |z|^2$.

(v) Now use the inequality from part (ii), the solution found in part (iii) [Remark: we found it using calculus with the assumption that the inner product was real, but it works for the complex case as well], and the lemma in part (iv) to prove the Cauchy-Schwartz inequality.

(b) Prove that $\|\cdot\|$ induced by an inner product $< \cdot, \cdot >$ is indeed a norm on V.

Hint: positive homogeneity and positive definiteness are easy. For the triangle inequality, try squaring both sides.

(c) Let $(V, \|\cdot\|)$ be a normed space. Define $d : V \times V \to \mathbb{R}$ by $d(x, y) = \|x - y\|$. Prove that (V, d) is a metric space.

Solution:

A. 1. Suppose $y = 0$. Then for some $z \in V$, $y = 0z$, so by linearity $\langle x, y \rangle = 0 \langle x, z \rangle = 0$, and since the function $\|\cdot\|$, as defined in Definition 20, is nonnegative, then the inequality must hold.

2. Recall that we are letting t take complex values in this calculation. Let

$$\begin{aligned} F_{x,y}(t) &= \langle x + ty, x + ty \rangle \\ &= \langle x, x \rangle + \bar{t} \langle x, y \rangle + t\overline{\langle y, x \rangle} + t\bar{t} \langle y, y \rangle \end{aligned}$$

We know from nonnegativity that $F_{x,y}(t) \geq 0$ for all t.

3. We assume for this calculation that we are dealing with a real inner product space and we only let t take real values (this makes it easy to use calculus), in which case $\overline{\langle x,y\rangle} = \langle x,y\rangle$. Then we have

$$F_{x,y}(t) = \|x\|^2 + 2t\langle x,y\rangle + t^2\|y\|^2,$$

and

$$F'_{x,y}(t) = 2\langle x,y\rangle + 2t\|y\|^2.$$

Setting this to 0 to find a critical point, we get

$$t_0 = \frac{\langle x,y\rangle}{\|y\|^2}.$$

Note that we needed to take care of the case $y = 0$ separately because the above value is undefined if $y = 0$.

4. Let $z \in \mathbb{C}$, with $z = x + iy$. Then we have $\bar{z} = x - iy$, so that

$$z\bar{z} = x^2 - i^2y^2 = x^2 + y^2 = |z|^2.$$

5. Returning to the complex case, but using the value of t_0 chosen above (which might be a complex number if the inner product is complex, but is nonetheless well-defined), we find that

$$\begin{aligned} F_{x,y}(t_0) &= \|x\|^2 + \bar{t}_0\langle x,y\rangle + t\overline{\langle x,y\rangle} + t_0^2\|y\|^2 \\ &= \|x\|^2 - \frac{\overline{\langle x,y\rangle}}{\|y\|^2}\langle x,y\rangle - \frac{\langle x,y\rangle}{\|y\|^2}\overline{\langle x,y\rangle} + \frac{\overline{\langle x,y\rangle}\langle x,y\rangle}{\|y\|^4}\|y\|^2 \\ &= \|x\| - 2\frac{|\langle x,y\rangle|^2}{\|y\|^2} + \frac{|\langle x,y\rangle|^2}{\|y\|^2} \\ &= \|x\| - \frac{|\langle x,y\rangle|^2}{\|y\|^2} \end{aligned}$$

Since we know this value is nonnegative, then by solving for $|\langle x,y\rangle|^2$ we find that

$$|\langle x,y\rangle|^2 \le \|x\|^2\|y\|^2,$$

and then by taking the square root of both sides we get the Cauchy-Schwartz inequality,

$$|\langle x,y\rangle| \le \|x\|\,\|y\|.$$

B.
- *Positive Definitiveness:* This follows from nonnegativity of the inner product.
- *Positive Homogeneity:* Let $x \in V$, and let $\alpha \in \mathbb{C}$. Then $\langle \alpha x, \alpha x\rangle = \alpha\bar{\alpha}\langle x,x\rangle$ by linearity in the first parameter and conjugate linearity in the second parameter. Since $\alpha\bar{\alpha} = |\alpha|^2$, by the definition of $\|\cdot\|$ we obtain $\|\alpha x\|^2 = |\alpha|^2\|x\|$, which means $\|\alpha x\| = |\alpha|\,\|x\|$.
- *Triangle Inequality:* Since norms are nonnegative, it will suffice to show that

$$\|x+y\|^2 \le (\|x\| + \|y\|)^2 = \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2.$$

Using the fact that $z + \bar{z}$ is real and less than or equal to $2|z|$ to go from the second line to the third line and the Cauchy-Schwartz inequality to go from the third to the fourth, we find that

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \overline{\langle x,y\rangle} + \langle y,y\rangle \\ &\le \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\ &\le \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

Thus, $\|x+y\| \le \|x\| + \|y\|$.

C. Positive definiteness is obvious from positive definiteness of the norm. Symmetry comes from the homogeneity, since

$$\|y - x\| = \|(-1)(x-y)\| = \|x - y\|.$$

The triangle inequality follows from the triangle inequality if the norm: Let $x, y, z \in V$. Then

$$\|x - y\| = \|x - z + z - y\| \le \|x - z\| + \|z - y\|.$$