As promised, I have written an itemized list of topics we've covered in Math 165 since Midterm 1. This list of topics (and the proportion of time we've spend on them since the beginning of the semester) will align with the problems you will see on the midterm exam. In studying for the midterm, please note that I consider the homework exercises and the everything I've covered in lecture to be the best source of practice (problems, proofs, etc). If you know how to approach each problem/exercise/proof, are able to work quickly and accurately, and understand the theory and methodology by which you have obtained a solution/proof, you should perform well on the exam. At the end of this document, I will also list some practice problems that you should do (and maybe treat like a homework assignment). While I might have you turn some in, they will not be due next week.

Definitions:

The following list enumerates all the definitions you need to know by heart. In particular, you should make sure to know all quantifiers involved in the definitions and the order in which they appear. Also, for each definition, you should be able to come up with several examples satisfying the definition (and hopefully things that don't satisfy the definition).

1. You should know the definition of the d-dimensional Euclidean space, $\mathbb{R}^d$. We shall refer to the elements $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$ as both *vectors* and *points* and you should be comfortable seeing them from both perspectives. We remark that the k-th coordinate of the vector $\mathbf{x} = (x_1, x_2, \ldots, x_d)$ is the real number (or scalar) x_k; this is the number appearing in the kth entry of $\mathbf{x}$. Finally, it will often be useful to write these vectors as "column vectors" and so write them as
$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}.$$

2. You should know how to add $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^d$ to form the vector $\mathbf{x} + \mathbf{y} \in \mathbb{R}^d$. Likewise, you should know how to subtract the vectors $\mathbf{x}$ and $\mathbf{y}$. Further, you should know how to scale a vector $\mathbf{x}$ by a scalar $\alpha \in \mathbb{R}$ to form a vector $\alpha\mathbf{x} \in \mathbb{R}^d$. You should also know what these operations mean geometrically, especially in $\mathbb{R}^2$ and $\mathbb{R}^3$. For example, for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, can you draw $\mathbf{x} \pm \mathbf{y}$ and $\alpha\mathbf{x}$ for various values of $\alpha \in \mathbb{R}$?

3. You should know that the set of vectors $\mathbb{R}^d$ satisfy the "vector space" axioms, This appears as Theorem 8.2 (in Wade Ed. 4)

4. You should know the definition of the dot product of vectors $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^d$. This is, by definition the scalar
$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_d y_d = \sum_{k=1}^{d} x_k y_k$$
for $\mathbf{x} = (x_1, x_2, \ldots, x_d)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_d) \in \mathbb{R}^d$.

5. You should know the definition of the Euclidean norm of a vector $\mathbf{x} \in \mathbb{R}^d$. This is, by definition, the non-negative quantity
$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_d^2}$$
defined for $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$. The Euclidean norm is sometimes called the ℓ^2 norm and written $\|\cdot\|_2$. Two other norms of interest, the ℓ^1 and ℓ^∞ norms, are given by
$$\|\mathbf{x}\|_1 = \sum_{k=1}^{d} |x_k|$$

end

$$\|\mathbf{x}\|_\infty = \max_{k=1,2,\dots,d} |x_k|,$$

respectively.

6. Armed with the Euclidean norm, $\|\cdot\|$, we can define the Euclidean distance between points $\mathbf{a}$ and $\mathbf{b} \in \mathbb{R}^d$ by

$$d(\mathbf{a}, \mathbf{b}) = \|\mathbf{b} - \mathbf{a}\|.$$

As we discussed in class, this is the distance yielded by the Pythagorean theorem, i.e., it's the "shortcut" distance. There are other distances associated with the ℓ^1 and ℓ^∞ norms in $\mathbb{R}^d$ and it might be a useful exercise to see what they measure in $\mathbb{R}^2$. For instance, one of them is referred to as the "taxicab" distance – can you tell which one?

7. Given vectors $\mathbf{a}$ and $\mathbf{b} \in \mathbb{R}^d$, you should know that the line the passes through $\mathbf{a}$ and $\mathbf{b}$ is the line

$$L = \{\mathbf{a} + t(\mathbf{b} - \mathbf{a}) \in \mathbb{R}^d : t \in \mathbb{R}\}.$$

8. Also, the line through the vector $\mathbf{a} \in \mathbb{R}^d$ in the direction of $\mathbf{v} \in \mathbb{R}^d$ is the set

$$L = \{\mathbf{a} + t\mathbf{v} \in \mathbb{R}^d : t \in \mathbb{R}\}.$$

9. Motivated by the fact that $\mathbf{x}\cdot\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|\cos(\theta)$ where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$ (this is true geometrically when $d = 2, 3$ as a consequence of the law of cosines and taken as a definition when $d > 3$), we say that vectors $\mathbf{x}$ and $\mathbf{y}$ are perpendicular or orthogonal when $\mathbf{x} \cdot \mathbf{y} = 0$. In this case, we write $\mathbf{x} \perp \mathbf{y}$.

10. A vector $\mathbf{x}$ is called unit if $\|\mathbf{x}\| = 1$

11. We say that a collection of vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ is orthonormal if each vector is a unit vector and $\mathbf{u}_k \perp \mathbf{u}_j$ whenever $j \neq k$. In other words,

$$\mathbf{u}_j \cdot \mathbf{u}_k = \begin{cases} 0 & j \neq k \\ 1 & j = k. \end{cases}$$

12. You should know the vectors we've referred to as the standard basis (or coordinate) vectors in $\mathbb{R}^d$. These are the vectors

$$\mathbf{e}_k = (0, 0, \dots, 0, 1, 0, \dots, 0)$$

where the 1 appears only in the kth coordinate. So, in $\mathbb{R}^2$ for example, $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$. In $\mathbb{R}^3$, these are also written $\mathbf{e}_1 = \mathbf{i}$, $\mathbf{e}_2 = \mathbf{j}$, and $\mathbf{e}_3 = \mathbf{k}$. Can you verify that, in any dimension d, $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_d$ form an orthonormal collection. In fact, they also form a so-called *orthonormal basis* in the sense that every element $\mathbf{x} \in \mathbb{R}^d$ can be written in the form

$$\mathbf{x} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_d\mathbf{e}_d$$

for some constants $a_1, a_2, \dots, a_d$. Can you determine these constants using the dot product and its properties?

13. You should know the definitions of the parallelograms and parallelepipeds given in Section 8.2 of Wade.

14. You should also know the definition of hyperplanes in $\mathbb{R}^d$. These are given as sets of points

$$P = \{\mathbf{x} \in \mathbb{R}^d : n_1x_1 + n_2x_2 + \cdots + n_dx_d = c\}$$

where $n_1, n_2, \dots, n_d$ and c are all constants. In fact, this can be written equivalently as

$$P = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{n} \cdot \mathbf{x} = c\}$$

where $c \in \mathbb{R}$ and $\mathbf{n} = (n_1, n_2, \dots, n_d) \in \mathbb{R}^d$ is a vector which is pependicular to the plane. There are several ways to produce such equations (some include the cross product in the context of $\mathbb{R}^3$ and you should be able to do many example. See the computation section below.

15. As some basic differentiability will be on this exam, and so linear/affine functions are very important, it is useful to real the material on linear transformations that appear in Section 8.2 of Wade's book.

16. Special to $\mathbb{R}^3$ is the cross product. That is, for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$,

$$\mathbf{x} \times \mathbf{y} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} = (x_2 y_3 - x_3 y_2, y_1 x_3 - y_3 x_1, x_1 y_2 - x_2 y_1) \in \mathbb{R}^3.$$

17. Given a sequence of points/vectors $\{\mathbf{x}_n\} \subseteq \mathbb{R}^d$, you should know what it means for these vectors to converge to another vector $\mathbf{x}$. This is the following definition:

    Definition 1. *A sequence of vectors $\{\mathbf{x}_n\} \subseteq \mathbb{R}^d$ converges to a vector $\mathbf{x} \in \mathbb{R}^d$ if, for every $\epsilon > 0$, there is some natural number N for which*

    $$\|\mathbf{x}_n - \mathbf{x}\| < \epsilon$$

    whenever $n \geq N$. In this case we write

    $$\mathbf{x} = \lim_{n \to \infty} \mathbf{x}_n.$$

18. You should know what it means for a sequence of points/vectors $\{\mathbf{x}_n\}$ to be bounded and Cauchy. Bounded means that, for some sufficiently large M,

    $$\|\mathbf{x}_n\| \leq M$$

    for all n. Equivalently, the entire sequence lives in some closed Euclidean ball $\overline{B}_M(0) = \{\mathbf{x} \in \mathbb{R}^d : \|x\| \leq M\}$ centered at the origin. Here is the definition of Cauchy:

    Definition 2. *A sequence $\{\mathbf{x}_n\} \subseteq \mathbb{R}^d$ is said to be Cauchy if, for all $\epsilon > 0$, there is some natural number N for which*

    $$\|\mathbf{x}_n - \mathbf{x}_m\| < \epsilon$$

    for all $n, m \geq N$.

19. You should know what it means for a subset C of $\mathbb{R}^d$ to be closed, i.e., it contains all of its limit points. In other words, if $\{\mathbf{x}_n\}$ is a convergent sequence of points, all of which are in C, then the sequence's limit is also a member of C. Also, you should know what it means for a subset K of $\mathbb{R}^d$ to be compact. For us, this means that the set is both closed and bounded.

20. You should know what it means for a subset $\mathcal{O}$ of $\mathbb{R}^d$ to be open. Just like with $\mathbb{R}$, open means that, for every $\mathbf{x} \in \mathcal{O}$, there is some ϵ-neighborhood of $\mathbf{x}$ that is completely containing in $\mathcal{O}$, i.e., there exists some $\epsilon > 0$ for which $B_\epsilon(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^d : \|\mathbf{x} - \mathbf{y}\| < \epsilon\} \subseteq \mathcal{O}$.

21. You should have a good idea of what a function F from $\mathbb{R}^n$ to $\mathbb{R}^m$ is. Precisely, given some set $\mathcal{D} \subseteq \mathbb{R}^n$ (called the domain of f), a function $F : \mathcal{D} \to \mathbb{R}^m$ is a function mapping points $\mathbf{x} \in \mathcal{D}$ to points $\mathbf{y} = F(\mathbf{x}) \in \mathbb{R}^m$. By an abuse of notation, we will often write $F : \mathbb{R}^n \to \mathbb{R}^m$ to mean a function $F : \mathcal{D} \to \mathbb{R}^m$ for $\mathcal{D} \in \mathbb{R}^n$ (i.e., the function need not be defined everywhere on $\mathbb{R}^n$). If we write the outputs as column vectors, these functions are necessarily of the form

    $$F(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix}$$

    where, for each $k = 1, 2, \ldots, m$, $f_k : \mathcal{D} \to \mathbb{R}$ is called the kth-coordinate function of F. Each coordinate function is a so-called scalar-valued function (meaning that its codomain is $\mathbb{R}$).

22. You should know the definition of the limit of a function:

Definition 3. *Let $F : \mathcal{D} \to \mathbb{R}^m$ where $\mathcal{D} \subseteq \mathbb{R}^n$ and let $\mathbf{x}_0$ be a limit point of $\mathcal{D}$, i.e., for every $\epsilon > 0$, the open ball $B_\epsilon(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x} - \mathbf{x}_0\| < \epsilon\}$ contains a point of $\mathcal{D}$ distinct from $\mathbf{x}_0$ (draw this, what does it mean?). Given a vector $\mathbf{y} \in \mathbb{R}^m$, we say that that function $F(\mathbf{x})$ converges to $\mathbf{y}$ as $\mathbf{x} \to \mathbf{x}_0$ if, for every $\epsilon > 0$, there is some $\delta > 0$ such that*

$$\|F(\mathbf{x}) - \mathbf{y}\| < \epsilon$$

whenever $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$. In this case, we write $\lim_{\mathbf{x} \to \mathbf{x}_0} F(\mathbf{x}) = \mathbf{y}$.

23. Please know the following definition of continuity.

Definition 4. *Let $F : \mathcal{D} \to \mathbb{R}^m$ for $\mathcal{D} \subseteq \mathbb{R}^n$. We say that F is continuous at $\mathbf{x}_0 \in \mathcal{D}$ if $\lim_{\mathbf{x} \to \mathbf{x}_0} F(\mathbf{x}) = F(\mathbf{x}_0)$. In other words, for every $\epsilon > 0$, there is a $\delta > 0$ for which*

$$\|F(\mathbf{x}) - F(\mathbf{x}_0)\| < \epsilon$$

whenever $\|\mathbf{x} - \mathbf{x}_0)\| < \delta$.

24. Know the following definition of partial derivative: Given a scalar-valued function $f : \mathcal{D} \to \mathbb{R}$ for $\mathcal{D} \subseteq \mathbb{R}^d$ and $\mathbf{x} \in \mathcal{D}$, for each $k = 1, 2, \ldots, d$, the partial derivative of f with respect to x_k at $\mathbf{x} = (x_1, x_2, \ldots, x_d)$ is given by

$$\frac{\partial f}{\partial x_k}(\mathbf{x}) = \lim_{h \to 0} \frac{f(x_1, x_2, \ldots, x_k + h, \ldots, x_d) - f(x_1, x_2, \ldots, x_d)}{h} = \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{e}_k) - f(\mathbf{x})}{h}$$

whenever this limit exists.

25. More general than the above, given a function $f : \mathcal{D} \to \mathbb{R}$ for $\mathcal{D} \subseteq \mathbb{R}^d$ and a vector $\mathbf{v} \in \mathbb{R}^d$ (where it is often demanded that $\mathbf{v}$ is a unit vector), the directional derivative of f in the direction $\mathbf{v}$ (or along $\mathbf{v}$) at the point $\mathbf{x} \in \mathcal{D}$ is defined by

$$D_{\mathbf{v}}(f)(\mathbf{x}) = \mathbf{v}(f)(\mathbf{x}) = \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{v}) - f(\mathbf{x})}{h}$$

provided this limit exists.

26. Our last definition covered on the exam is that of differentiability of a function. This is

Definition 5. *Let $F : \mathcal{D} \to \mathbb{R}^m$ where $\mathcal{D} \subseteq \mathbb{R}^n$ and suppose that $\mathbf{x}$ is an interior point of $\mathcal{D}$. We say that F is differentiable at $\mathbf{x}_0$ if there exists a $m \times n$ matrix A for which*

$$F(\mathbf{x}) = F(\mathbf{x}_0) + A(\mathbf{x} - \mathbf{x}_0) + \|\mathbf{x} - \mathbf{x}_0\|\mathcal{E}(\mathbf{x} - \mathbf{x}_0)$$

where $\mathcal{E} : \mathbb{R}^n \to \mathbb{R}^m$ is function which is continuous at $\mathbf{0} \in \mathbb{R}^n$ with $\mathcal{E}(\mathbf{0}) = \mathbf{0}$. In this case, we say that A is the derivative matrix (or total derivative) of F at $\mathbf{x}_0$ and write $D(F)(\mathbf{x}_0) = A$.

We make two remarks about this definition, both of which will be elucidated on the Monday following Spring break.

(a) The application of the matrix A to the vector $\mathbf{x} - \mathbf{x}_0$ must be done with $\mathbf{x} - \mathbf{x}_0$ written as column vectors in $\mathbb{R}^n$. Precisely, for $\mathbf{h} = \mathbf{x} - \mathbf{x}_0 \in \mathbb{R}^n$, this is

$$A(\mathbf{x} - \mathbf{x}_0) = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} a_{11}h_1 + a_{12}h_2 + \cdots + a_{1n}h_n \\ a_{21}h_1 + a_{22}h_2 + \cdots + a_{2n}h_n \\ \vdots \\ a_{m1}h_1 + a_{m2}h_2 + \cdots a_{mn}h_n \end{pmatrix}$$

(b) Consistent to our one-variable theory, the above definition says that F has a best affine approximation at $\mathbf{x}_0$. In this case, this approximation is $A(\mathbf{x} - \mathbf{x}_0) + F(\mathbf{x}_0)$.

Results (Theorems, propositions, lemmas, corollaries):

You should know all of the results we discussed (and especially those we proved) in class. I'll list most of these results. Unless otherwise mentioned, you should know the statement of the result precisely and have a really good idea of how they are proved – ideally, you should be able to reproduce the proof. Still, I don't feel it makes sense for you to spend all of your time memorizing the proofs of these results. It certainly makes sense to spend a lot of time deeply understanding the proofs, but you should spend equal time working examples (including creating your own), counter examples, and solving problems. For example, can you cook up and work through an example to illustrate every result in the list below? For certain results, can you create an example for which the conclusion fails when the hypotheses aren't quite met?

1. You should have a good idea of the vector space properties for $\mathbb{R}^d$ and have a strong idea of what's going on geometrically. I believe this is the Theorem 8.2 in the 4th edition of Wade.

2. You should know the algebraic properties of the dot product. Some of these appear in Theorem 8.2 in the 4th edition of Wade.

3. The Cauchy-Schwarz inequality. Also, the fact that the ratio

$$\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\in[-1,1]$$

allows us to find an angle θ for which

$$\mathbf{x}\cdot\mathbf{y}=\|\mathbf{x}\|\|\mathbf{y}\|\cos(\theta).$$

Of course, you should also understand that this angle θ is precisely that which is given by the law of cosines in $\mathbb{R}^2$ and $\mathbb{R}^3$ (as we saw in class). Thus in dimensions 2 and 3, it is the angle between the vectors $\mathbf{x}$ and $\mathbf{y}$. In $\mathbb{R}^d$ for $d>3$, we take this as a definition of the angle between vectors even if we do not have a definition for such a thing, *a priori*. This is a standard way to assign geometric meaning to a space that, perhaps, you cannot see. On this note, you should watch the video this video by the late (and great) astronomer and science popularizer, Carl Sagan.

4. You should know how the triangle inequality follows from the Cauchy-Schwarz inequality.

5. Here is a nice example: For a given vector $\mathbf{x}=(x_1,x_2,\dots,x_d)\in\mathbb{R}^d$, consider a vector $|\mathbf{x}|=(|x_1|,|x_2|,\dots,|x_d|)\in\mathbb{R}^d$ and another vector $\mathbf{1}=(1,1,1,1,\cdots,1)$. Use the Cauchy-Schwarz inequality to prove that

$$\frac{1}{\sqrt{d}}\|\mathbf{x}\|_1\leq\|\mathbf{x}\|$$

where the ℓ^1 norm on the left can be seen as a dot product of $|\mathbf{x}|$ with $\mathbf{1}$; here, $\|\mathbf{x}\|$ means the regular Euclidean (ℓ^2) norm on the right. You can combine this result with the ones you proved in homework to see that

$$\frac{1}{\sqrt{d}}\|\mathbf{x}\|_1\leq\|\mathbf{x}\|\leq\sqrt{d}\|\mathbf{x}\|_\infty.$$

for every $\|\mathbf{x}\in\mathbb{R}^d$. This inequality turns out to be quite useful in many applications of analysis, vector calculus, and linear programming. It can also give a hint (look at how it turns useless as $d\to\infty$) to the problems that arise in infinite dimensions.

6. You should be able to prove that, for any vector $\mathbf{x}\in\mathbb{R}^d$,

$$\mathbf{x}=x_1\mathbf{e}_1+x_2\mathbf{e}_2+\cdots+x_d\mathbf{e}_d.$$

In fact, the following is true (some of it will be realized only when you take MA253): If $\{\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_d\}$ is an orthonormal collection of d vectors in $\mathbb{R}^d$, then any vector $\mathbf{x}\in\mathbb{R}^d$ can be expressed as

$$\mathbf{x}=a_1\mathbf{u}_1+a_2\mathbf{u}_2+\cdots a_d\mathbf{u}_d$$

where

$$a_k = \mathbf{u}_k \cdot \mathbf{x}$$

for $k = 1, 2, \ldots, d$. See if you can see why the coefficients a_k must be given by this dot product. In fact, this idea was key to the way we defined Fourier coefficients (in that context, we were using a different kind of "dot" product and were living in infinite dimensions).

7. You should know the basic algebraic properties of the cross product for vectors in $\mathbb{R}^3$.

8. You should also know the geometric properties of the cross product. For example, if $\mathbf{x}$ and $\mathbf{y}$ are non-zero vectors in $\mathbb{R}^3$, $\mathbf{x}\times\mathbf{y}$ is orthogonal to both $\mathbf{x}$ and $\mathbf{y}$. You should also know the right-hand rule and the following fact: For $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^3$,

$$\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\|\mathbf{y}\| \sin(\theta)$$

where θ is the angle (with $0 \leq \theta \leq \pi$) between $\mathbf{x}$ and $\mathbf{y}$.

9. Areas and volumes of parallelograms/parallelepipeds via cross products and determinants.

10. Prove that a sequence converges if and only if it is Cauchy. You should be able to do this as you did in your homwork. For completeness, here is a sketch of the argument:

Proof. Let $\{\mathbf{x}_n\}$ be a sequence in $\mathbb{R}^d$. If $\{\mathbf{x}_n\}$ is convergent, say to $\mathbf{x}$, then for any $\epsilon > 0$, we may choose an N for which

$$\|\mathbf{x}_n - \mathbf{x}\| < \frac{\epsilon}{2}$$

whenever $n \geq N$. Thus, by the triangle inequality, we have

$$\|\mathbf{x}_n - \mathbf{x}_m\| = \mathbf{x}_n - \mathbf{x} + \mathbf{x} - \mathbf{x}_n\| \leq \|\mathbf{x}_n - \mathbf{x}\| + \|\mathbf{x} - \mathbf{x}_m\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for $n \geq N$. Hence $\{\mathbf{x}_n\}$ is Cauchy.

Conversely, let's assume that $\{\mathbf{x}_n\}$ is Cauchy. Observe that, for each k, we have

$$|x_n k - x_m k| \leq \|\mathbf{x}_n - \mathbf{x}_m\|$$

for every $n, m \in \mathbb{N}$; here $x_n k$ denotes the kth coordinate of the vector $\mathbf{x}_n$ (I believe Wade writes this as $x_n(k)$ which is fine notation too). In any case, our assumption that $\mathbf{x}_n$ is Cauchy paired with the above inequality gives us that the sequence of numbers $\{x_n k\}$ is Cauchy. By the completeness of the real numbers, this sequence is convergent to some number which we will call x_k, i.e., $\lim_{n\to\infty} x_{nk} = x_k$. We can do this for every k thus producing a collection of number $x_1, x_2, \ldots, x_d$ to which our sequences of coordinates converge. With these, we may set

$$\mathbf{x} = (x_1, x_2, \ldots, x_d)$$

and we claim that, in fact, our original sequence converges to this $\mathbf{x}$. To see this, we select $\epsilon > 0$ and choose natural numbers $N_1, N_2, \ldots, N_d$ for which

$$|x_{nk} - x_n| < \frac{\epsilon}{\sqrt{d}}$$

for $n \geq N_k$ for each $k = 1, 2, \ldots, d$. Taking N to be the maximum of $N_1, N_2, \ldots, N_d$, we have

$$\|\mathbf{x}_n - \mathbf{x}\| \leq \sqrt{d}\|\mathbf{x}_n - \mathbf{x}\|_\infty = \max_{k=1}^{d} \sqrt{d}|x_{nk} - x_k| < \sqrt{d}\frac{\epsilon}{\sqrt{d}} = \epsilon$$

whenever $n \geq N$; here, we have used the relationship between the ℓ^2 and ℓ^∞ norms we discussed previously. Thus, $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$. □

You should also understand the argument that also gives this using the Bolzano-Weierstrass theorem.

11. We notice that our proof above relied on one-dimensional results. Leo Livshits calls this technique "back to the USSℝ" inspired by this song by the Beatles. Use the same idea to prove the Bolzano-Weierstrass theorem in $\mathbb{R}^d$:

    Theorem 6. *Let $\{\mathbf{x}_n\}$ be a bounded sequence in $\mathbb{R}^d$. Then $\{\mathbf{x}_n\}$ has a convergent subsequence.*

    Hint: First show that the first coordinate is a bounded sequence and so you can find a subsequence of the first coordinate that converges by the methods of one-dimension. Restricting the full vectors $\mathbf{x}_n$ to the indices n_k for which the first coordinate sequence converges, we have another bounded sequence. So, the second sequence is now bounded and we can continue this way. Each time, we continue to a new subsequence, which is a further subsequence of the previous one, until we get all coordinates to converge.

12. Along the lines of the above, can you prove that a sequence $\{\mathbf{x}_n\} \subseteq \mathbb{R}^d$ converges if and only if all of its coordinates converge?

13. You should be able to prove the following theorem about the convergence of vector-valued functions:

    Theorem 7. *Let $\mathcal{D} \subseteq \mathbb{R}^n$, $F : \mathcal{D} \to \mathbb{R}^m$ and take $\mathbf{x}_0$ to be a limit point of $\mathcal{D}$. Then*

    $$\lim_{\mathbf{x}\to\mathbf{x}_0} F(\mathbf{x}) = \mathbf{y}$$

    if and only if, for every sequence $\{\mathbf{x}_n\} \subseteq \mathcal{D}$ with $\mathbf{x}_n \to \mathbf{x}_0$, we have $F(\mathbf{x}_n) \to \mathbf{y}$.

14. Can you use the previous item to show that the composition of continuous functions is continuous? You should be able to.

15. You should know the extreme value theorem:

    Theorem 8. *Let K be a compact subset of $\mathbb{R}^d$ and $f : K \to \mathbb{R}$ a continuous function. Then $f(K)$ is bounded and there exist points $\mathbf{x}_m$ and $\mathbf{x}_M$ in K for which*

    $$f(\mathbf{x}_m) = \inf f(K) \qquad \text{and} \qquad f(\mathbf{x}_M) = \sup f(K).$$

    You should also have a very good idea of its proof, which we did in class.

16. You should know the following version of the intermediate value theorem:

    Theorem 9. *Let $\mathcal{D}$ be a connected subset of $\mathbb{R}^d$, let $f : \mathcal{D} \to \mathbb{R}$ and let $\mathbf{a}, \mathbf{b} \in C$. Then, for any real number c between $f(\mathbf{a})$ and $f(\mathbf{b})$, there is a point $\mathbf{c} \in \mathcal{D}$ for which $f(\mathbf{c}) = c$.*

    In class, we proved a slightly weaker statement where we asked only that $\mathcal{D}$ be path connected (which is a little more restrictive than $\mathcal{D}$ just being connected). In any case, you should understand the proof in the path connected case. You should also understand which types of sets are path connected and be able to investigate path connectedness of specific sets (like the ones on your homework).

17. Though you will really only be responsible for the definition (and idea/concept) of differentiability on the second midterm, you should be very comfortable with the definition and be able to verify things like: If a function $F : \mathcal{D} \to \mathbb{R}^n$ is differentiable at a point $\mathbf{x}_0 \in \mathcal{D} \subseteq \mathbb{R}^n$, then it is continuous at $\mathbf{x}_0$.

0.1 Things you should be able to do:

1. You should be able to do many examples and computations with vectors.
2. Given a vector $\mathbf{x}_0$ and $\mathbf{n} \in \mathbb{R}^d$, can you write down the equation of a plane through $\mathbf{x}_0$ and perpendicular to $\mathbf{n}$?
3. If I gave you three vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ in $\mathbb{R}^3$, could you find the equation of a plane through those vectors/points? Hint: $\mathbf{b} - \mathbf{a}$ and $\mathbf{c} - \mathbf{a}$ must live in the plane and so $(\mathbf{b} - \mathbf{a}) \times (\mathbf{c} - \mathbf{a})$ must perpendicular to the plane ....
4. You should be able to find volumes of parallelograms and parallelepipeds.
5. Given a set $A \subseteq \mathbb{R}^d$, you should be able to tell if the set if open/closed/path-connected/compact.
6. You should be able to determine if a given sequence $\{\mathbf{x}_n\}$ is
   (a) convergent
   (b) bounded
   (c) Cauchy

   and verify (using definitions/theorems rigorously) your assertions. If a given sequence is bounded, could you extract a convergent subsequence?
7. Given a function $F : \mathbb{R}^n \to \mathbb{R}^m$ (especially when $m = 1$) and a point $\mathbf{x}_0$, you should be able to determine if the limit
   $$\lim_{\mathbf{x} \to \mathbf{x}_0} F(\mathbf{x})$$
   exists. You should also be able to prove your assertion. On this, it would be useful to work through of the examples in Section 9.3 of Wade Ed. 4.
8. You should be able to determine if/when partial derivatives of scalar-valued functions exist. You should be able to compute them.
9. You should be able to determine if/when directional derivatives of scalar-valued functions exist. You should be able to compute them.
10. You should be able to determine if/when a function $F : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a point $\mathbf{x}_0$ and (if it is) identify its derivative matrix $D(F)(\mathbf{x}_0)$ at the point.

Things not to worry about:

You know about the following things and why they are important (at least for us) but I won't ask you anything detailed about them.

1. You don't need to worry about anything in Chapter 10 (on Metric spaces) of Wade. You will learn this material in MA338.
2. I did skip around Chapters 8 and 9 a bit and so there are many results in there I'm not making you responsible for. I'd stick to my list above. If you're curious if something will be covered, just email me and I'll tell you.

Some exercises

Exercise 1. 1. Find the equation of the plane through $(1, 2, -1, 1)$ and perpendicular to $(1, 0, 0, 1)$ in $\mathbb{R}^4$.

2. Find the equation of the plane through the points $(1, 1, 1)$, $(1, -1, 0)$, and $(2, 0, 1)$.

3. Find the volume of the parallelepiped spanned by the three vectors above.

Now, make up your own examples and do them.

Exercise 2 (Point to Plane Distance). In this exercise, we develop two methods of determining the distance between a point and a plane in $\mathbb{R}^3$. First, let's do a one-dimensional problem that you can solve easily.

1. Let $\mathbf{p} = (p_1, p_2)$ be a point in $\mathbb{R}^2$ and consider a line whose equation is $ax + by = c$ for some constants $a, b, c \in \mathbb{R}$. The distance between the point $\mathbf{p}$ and a point (x, y) on the line is

$$\text{dist} = \sqrt{(x - p_1)^2 + (y - p_2)^2}$$

In the case that $b \neq 0$, we can solve the equation $ax + by = c$ to find that $y = (c - ax)/b$ and so

$$\begin{aligned} \text{dist} &= \sqrt{(x - p_1)^2 + ((c - ax)/b - p_2)^2} \\ &= \sqrt{(x - p_1)^2 + \frac{1}{b^2}(c - ax - bp_2)^2} \end{aligned}$$

This can now be seen as a single-variable calculus problem. If we want to minimize the distance, it suffices to simply minimize the square of the distance and so we seek a point $x_0 \in \mathbb{R}$ for which

$$0 = \frac{d}{dx}\text{dist}^2 = \frac{d}{dx}\left((x - p_1)^2 + \frac{1}{b^2}(c - ax - bp_2)^2\right).$$

Find this x_0, use the equation of the plane to find the corresponding y_0, and plugging this back into the original equation will give you the minimum distance. Compute that too.

2. The answer that you obtained in the previous item can be done using some very simple trigonometry. If you consider the vector in $\mathbf{v} \in \mathbb{R}^2$ going from $\mathbf{p} = (p_1, p_2)$ to a point on the line (x, y), we have $\mathbf{v} = (x - p_1, y - p_2)$. If we you draw the line and the vector $\mathbf{v}$, the minimal distance from the point $\mathbf{p}$ to the plane can be seen as the length of the adjacent side of a right triangle with hypotenuse $\mathbf{v}$. The direction of this adjacent side is precisely given by a vector normal to the line (which, in this case, is the vector (a, b)). Consequently, the distance is

$$\text{dist} = \|\mathbf{v}\| \cos(\theta)$$

where θ is the angle between the vector $\mathbf{v}$ and the normal vector (a, b) (or its negative). If we can find a unit vector $\mathbf{u}$ which is parallel to (a, b), we then have

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos(\theta) = 1\|\mathbf{v}\| \cos(\theta) = \text{dist}.$$

To make sure that no negative sign is introduced, we have

$$\text{dist} = |\mathbf{u} \cdot \mathbf{v}|.$$

Now, it is easy to see that

$$\mathbf{u} = \frac{(a, b)}{\sqrt{a^2 + b^2}} = \left(\frac{a}{\sqrt{a^2 + b^2}}, \frac{b}{\sqrt{a^2 + b^2}}\right)$$

is a unit vector in the direction[1] of (a, b). So, note that $(x, y) = (0, c/b)$ is on the line so that $\mathbf{v} = (-p_1, c/b - p_2)$ is a vector from the point $\mathbf{p} = (p_1, p_2)$ to the plane. Okay, now compute the distance. Is it what you had before?

[1] Also, if I gave you any non-zero vector $\mathbf{x}$, could you find a unit vector in its direction?

3. In the previous item, you actually used something called a projection (which we haven't discussed in class). Given a non-zero vector $\mathbf{n} \in \mathbb{R}^d$ and a vector $\mathbf{v} \in \mathbb{R}^d$, the vector and scalar projections of $\mathbf{v}$ onto $\mathbf{n}$ are defined, respectively, by
$$\mathbf{Proj_n}(\mathbf{v}) = \frac{\mathbf{n}\cdot\mathbf{v}}{\|\mathbf{n}\|^2}\mathbf{n} \qquad \text{and} \qquad \mathrm{Proj}_{\mathbf{n}}(\mathbf{v}) = \|\mathbf{Proj_n}(\mathbf{v})\| = \frac{\mathbf{n}\cdot\mathbf{v}}{\|\mathbf{n}\|}.$$
Note that the projection vector is certainly parallel to $\mathbf{n}$. Also note, in the previous item $\mathbf{n} = (a, b)$.

4. Okay, use the above to find the point-to-plane distance in $\mathbb{R}^d$. Notice that, for a plane with the equation
$$ax + by + cz = d$$
$\mathbf{n} = (a, b, c)$ is a normal vector to the plane (you should know this well). So, if we take a point $\mathbf{p} = (p_1, p_2, p_3) \in \mathbb{R}^3$, we should be able to produce a point to plane distance via a projection. Do it.

5. In particular, compute the point to plane distance between $(2, 1, 1)$ and the plane $x + y + z = 1$. Does your answer make sense?

6. Okay, now lets find the point to plane distance using multivariable calculus and, in particular, partial derivatives. The distance between $(2, 1, 1)$ and a point (x, y, z) on the plane $x + y + z = 1$ is
$$\text{dist} = \sqrt{(x-2)^2 + (y-1)^2 + (z-1)^2}$$
We can remove one variable to see that $z = 1 - x - y$ so that
$$\text{dist} = \sqrt{(x-2)^2 + (y-1)^2 + (-x-y)^2} = \sqrt{(x-2)^2 + (x+y)^2 + (y-1)^2}.$$
Of course, to minimize the distance, we can minimize its square and so we'll consider
$$f(x, y) = (x-2)^2 + (x+y)^2 + (y-1)^2.$$
Find (x_0, y_0) for which
$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \qquad \text{and} \qquad \frac{\partial f}{\partial y}(x_0, y_0) = 0$$
Does the distance yielded by this minimization give you the answer using projections?

7. Only do this if you are ambitious: Do the following example in general. So $\mathbf{p} = (p_1, p_2, p_3)$ is a point and $ax + by + cz = d$ is the equation of a plane. Minimize a function $f(x, y)$ by finding points where its partial derivatives are 0 to produced a minimum.

Exercise 3. Okay, here are some exercises from Wade Ed. 4.

1. Do Exercises 8.2.1-8.2.4 of Wade Ed. 4.

2. Do Exercise 9.1.1-9.1.4 of Wade Ed. 4.

3. Do Exercises 9.3.1-9.3.4 of Wade Ed. 4.

Exercise 4. Here are some examples as promised. The first is a simpler version (that still has all salient features) than that discussed in class on Friday. The second is motivated by a fruitful discussion I had with our own David Chamberlain.

1. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by
$$f(x, y) = \begin{cases} 1 & y = x^2 \text{ for } x > 0 \\ 0 & \text{else} \end{cases}$$

(a) Prove that, for every path $\mathbf{r}(t) = t\mathbf{v} = (tv_1, tv_2)$, going to the origin as $t \to 0$,

$$\lim_{t\to 0} f \circ \mathbf{r}(t) = f(0,0) = 0$$

and hence f has a limit along all lines to $(0,0)$ and they agree with $f(0,0)$.

(b) Prove that, for any $\mathbf{v} \in \mathbb{R}^2$, the directional derivative $D_{\mathbf{v}}(f)(0,0) = \mathbf{v}(f)(0,0)$ exists and is zero at $(0,0)$.

(c) Prove that f is discontinuous.

2. Consider the function $g : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$g(x,y) = \begin{cases} x+y & xy > 0 \\ 0 & xy \le 0 \end{cases}$$

(a) Prove that g is continuous at $(0,0)$.

(b) Prove that, for any $\mathbf{v} \in \mathbb{R}^2$, the directional derivative $D_{\mathbf{v}}(g)(0,0)$ exists.

(c) Prove that g is not differentiable at $(0,0)$.

3. Consider the function $k : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$k(x,y) = \begin{cases} xy & xy > 0 \\ 0 & xy \le 0 \end{cases}$$

(a) Prove that k is continuous at $(0,0)$.

(b) Prove that, for any $\mathbf{v} \in \mathbb{R}^2$, the directional derivative $D_{\mathbf{v}}(k)(0,0)$ exists and is zero.

(c) Prove that k is differentiable at $(0,0)$.