#h1 Chapter 19 - Manifold Learning and Basis Functions

#h2 Manifolds as the superclass of machine learning

#h3 19.1 - Introduction

#pg Manifolds form an interesting combination of the algorithmic objectives we have previously discussed: regression, classification, clustering, and dimensionality reduction methods. In some sense, one can argue the latter are merely more basic subclasses with manifold learning being the superclass. Although this offers an interesting take on manifold learning theoretically, in practice our objective for some particular use case is usually narrower, such as classifying whether a tumor on a medical image is benign or malignant. As such, we will not dedicate a lot of this chapter on implementing manifold learning, but instead focus on the underlying conceptual connection between manifolds and other algorithm objectives.

#im ../assets/figures/019/019-01.gif 50 256 Figure 19.1 - Moebius Strip Manifold.

#br

#pg The idea of manifolds (in combination with the concept of basis functions we will touch on later in this chapter) are a perfect example of understanding machine learning in a geometric context. Modern deep belief neural networks (discussed in later chapters) are very large models, having millions if not billions of parameters needed to reach a particular level of performance. But it is often hard to say anything meaningful about what exactly it has learned. And, at least for me, this is where geometry come in, and allows us to visualize and describe the meaningful transformations from inputs to outputs.

#h3 19.2 - Color perception as an example of a true manifold

#pg We are aware that this is a book on data science, AI, and machine learning. But neuroscientific and perceptual psychology tackle highly similar theoretical questions and have in many instances provided the inspiration on how to build artificial systems with similar capabilities. While this appears to be mostly true in the perceptual realm, there is a shared understanding across many fields that study the brain or draw inspiration from it. By and large, brains run on general purpose machinery and computational principles. This is especially true for the part of the brain most developed in humans - the cortex. In this light, we can think of the brain of implementing its own machine learning algorithms, transforming inputs into suitable outputs. Furthermore, the models share a similar implementations and computations, and are mostly independent from the type of input data they operate on (retinal inputs, touch, sound). To explain manifolds, and their relationship to other algorithmic objectives we already encountered (e.g., classification, regression), we could use a data set in which the dimensions represent some interesting feature we might encounter in the data here at MSK. Interesting or not, combine a few of those features together, it is becoming a little hard to imagine the underlying structure of the data. We therefore decided to continue with the example we introduced in the last chapter: color. In the last chapter we focused on colors as they are produced by a screen or monitor, using the three primary colors red, green, and blue. Here, we will take it one step further, and works towards an understanding of the relationship between the light entering our eyes and our perception of color. In doing so, we will naturally work towards an understanding of what a manifold is, what it represents, and how to think of manifolds in relation to the data you are keen on exploring and using for to build a model. We would also like this opportunity to apologize to those reading this who do not have normal color vision. If so, please skip ahead to the example on human faces.

#im ../assets/figures/019/019-02.png 50 256 Figure 19.2 - I loved fingerpainting sessions in kindergarten. Given my tendency for getting most of the paint on my person and clothing, rather than the page, my parents were less enthusiastic.

#h3 19.3 - Newtown's Mistake

#pg A widely held belief is that Sir Isaac Newton proved that white light is composed of all colors, which he demonstrated by refracting sunlight through a prism and observing a rainbow-like pattern emerge on projection screen. Refraction refers to the fact that wavelengths that pass through a medium (such as glass) do so at different angles. Therefore, when a beam of light from the sun, which is composed of a large range of wavelengths, is passed through a prism, it has the effect of spreading out the light at different wavelengths, from shorter to longer wavelengths. And what Newtown observed is that each wavelength, light seems to have a different color, with the color varying smoothly with wavelength.

#im ../assets/figures/019/019-03.png 50 256 Figure 19.3 - Newton using a prism to refract sunlight from outside onto a white surface. Why this appears to cause quite a bit of alarm in the gentleman in the background is not known to me.

#br

#pg I think many modern textbooks still use this example to explain the relationship between light and color, but it is (and I hate to disagree with Sir Isaac Newton) wrong, at least when it comes our perception of color. The central flaw in Newtown's observation is that he essentially equated color with wavelength. One wavelength = one color. This is by no means how the visual system works. Yes, a laser at a particular frequency will appear red, and one at a different frequency will appear green. But our world does not exist of objects that emit single, narrowly defined wavelength. Instead, what you see is a smoothly varying distribution of wavelengths that your brain will need to interpret somehow as a specific color. In addition, while Newtown was able to obtain a rainbow's worth of colors out of his sunlight experiment, most people will notice that on closer examination many colors appear to be missing. As Goethe and Huygens, Sir Newton's contemporaries, aptly pointed out: something is either wrong or seriously lacking.

#im ../assets/figures/019/019-04.png 50 256 Figure 19.4 - The sensitivity of the three cone photoreceptors (short, medium, and long) in normal human trichromatic vision relative to the visual spectrum. L, M and S stand for long, medium, and short wavelength receptors, respectively. However, quite often they are referred to the red, green, and blue receptors, which is a misleading characterization. In particular, the sensitivity of each receptor type varies smoothly across a wide range of wavelengths. In addition, there is substantial overlap, especially between the medium and long wavelength receptors.

#h3 19.4 - What you probably have been taught incorrectly as a result

#pg So then how is color perceived and encoded? It all starts with the cone receptors in your eye's retina. In normal human trichromatic vision, your retina contains three different photoreceptors (men are more genetically prone to colorblindness which typically means they only have 2 out of 3, women on the other hand do occasionally have a fourth that varies slightly from the other three). These photoreceptors are maximally sensitive to different wavelengths of the spectrum. Therefore, often, you will read or hear people refer to the 'red', 'green’, and 'blue' photoreceptors. As you can see in Figure 19.4, there seems to be some truth to this: it looks like the peak sensitivity of each cone corresponds to these colors. However, this simplification in which our retina is like a computer screen in that it is built around using three primary colors is misleading. On closer inspection, the three cone types have dramatically broad and overlapping sensitivities. This means that they respond to a very broad range of different wavelengths and that different cone types more of less respond to the same wavelengths. This is especially true for the so called 'red' and 'green' cones.

#h3 19.5 - The dimensions of color vision - luminance

#pg You might think of this is a design flaw, but it is not. Far from it. It is ingenious. Because these cones are responsible not just for detecting what most people would call 'color', they also need to signal brightness of incoming light. Compare the renderings in Figure 19.4 in which the position of the sun was varied to mimic a certain time of the day. As the day progresses, the sun rises to the point where it shines directly into the room, and as a result it the room looks brighter. Brightness (also sometimes called lightness) is subjective experience of the overall amount of light, and its objective counterpart is referred to as luminance and is expressed in candela per square meter. And 1 candela is the luminance produced by a single average candle.

#im ../assets/figures/019/019-05.png 50 384 Figure 19.5 - Four renderings of the same 3D model of a living room, but under different lighting conditions, simulating the position of the sun as the day progresses. In colloquial language, we speak about brightness, which is the subjective experience of an increase in overall light. In contrast, luminance is the objective measurement standard, expressed in candela per square meter.

#h3 19.6 - The dimensions of color vision - saturation

#pg In addition, the cones also signal the vibrance of colors. Consider these two pieces of art in Figure 19.6. On the top is a mural by Dos Gemeos (Dos Gemeos means 'twins' which you can take literally: they are twin brother mural artists from Brazil). Some of you might remember that this mural was done on a still standing wall of a demolished building at the corner of the Bowery and Houston Street.

#im ../assets/figures/019/019-06.png 50 512 Figure 19.6 - Top panel: A Dos Gemeos mural, once on display on a wall at the intersection of the Bowery and Houston Street, New York City. Bottom panel: the Sleeping Gypsy by Henri Rousseau, part of the permanent collection of the MoMa, New York City. Both pieces of art use a wide variety of different colors. Yet, the Dos Gemeos mural appears to be much more vibrant. This physical counterpart of perceived vibrance is referred to as saturation.

#br

#pg On the bottom you can see one of the most famous paintings of the impressionist period, by the hand of Henri Rousseau: the sleeping Gypsy (on display at the MoMA). Provided you have normal color vision and are looking at these images under reasonably normal light conditions, you can see how both paintings use a rich palette of colors. But still there is a clear overall perceptual difference in color. The Dos Gemeos mural appears far more vibrant with its colors, whereas the Sleeping Gypsy leans more towards using almost pastel like colors. This difference in vibrance is referred to as saturation. The more a color is saturated, the more vibrant it appears, and low levels of saturation make it appear more muted and grayish.

#im ../assets/figures/019/019-07.png 50 160 Figure 19.7 - Two patches, one of them purple, one of them green

#br

#im ../assets/figures/019/019-08.png 50 160 Figure 19.8 - Correcting the green patch to match the luminance of the purple patch

#br

#im ../assets/figures/019/019-09.png 50 160 Figure 19.9 - Correcting for both luminance and saturation leaves only difference a in hue

#br

#pg These two examples allow to start unpacking what it means to see color beyond Sir Isaac's simplification and isolate it from other perceptual aspects that confound our understanding, systematically. Let's start with the two patches side by side in Figure 19.7. Assuming you have normal color vision, they are likely to differ to you in what most people would call their color. One of them is purple, and the other is green. But on closer inspection, they differ beyond simply their labels 'purple' and 'green". First, the purple appears much brighter, compared to the green. As if it was illuminated directly by the sun, whereas the green might be in the shadow of another object. As we shall see later, we can correct for this difference. So, in the next panel, the two patches have been made what is called isoluminant: we adjusted the green to the point where its luminance matches the purple. However, having done so, we now notice an additional difference: that of saturation, see Figure 19.8. So, once more we correct for this, by increasing the saturation of the green while making sure we do keep the two patches isoluminant.

#im ../assets/figures/019/019-10.png 50 192 Figure 19.10 - Starting from the same red on the left-hand side, depicted is a stepwise reduction in luminance (top), saturation (middle) and change in hue (bottom).

#h3 19.7 - The dimensions of color vision - hue

#pg Finally, we arrive at Figure 19.9. These two patches have the same luminance and the same amount of saturation. Yet they still differ. One of them is still a red, and one of them can still be called a green. This final perceptual difference is a difference in what is called hue (sometimes also referred to as chroma). Color, therefore, can be seen as the combination of these three dimensions of variation: color = lightness + saturation + hue. All three rows in Figure 19.10 start with the same-colored patch, but in each row, we varied only the luminance, saturation, or hue respectively. Selectively removing both any difference in lightness and saturation, setting the to some constant value, creates a so-called isoluminant image, which we did for the painting in Figure 19.11. As you can see, the detail of this painting clearly isn't just in variations in hue. Lightness and saturation are equally necessary to give the painting its vivid appearance.

#h3 19.8 - Representing color on a screen

#pg We realize that at this point, you might be even more confused about what this must do be machine learning, and how it relates to regression, classification, clustering, and dimensionality reduction. Trust us, we are almost there. The final

intuition is that the three-dimensional representation of color that humans use (luminance, saturation, and hue) is different from how we have been designing monitors.

#im ../assets/figures/019/019-11.png 50 384 Figure 19.11 - Giorgio de Dhirico' Piazza d'italia in its original format (top left), only showing luminance (top right), or saturation (bottom left, brighter gray indicate higher levels of saturation) and finally the painting where luminance has been fixed, rendering the painting as isoluminant. The only visible variation in this top right image is hue.

#br

#pg As we discussed in the previous chapter, the way that color is represented on a screen (or any light emitting device) is by using additive mixing of three primary colors: red, green, and blue. Therefore, A color on a screen can be represented as a 3-dimensional vector of these three colors. For example, red is represented as [1,0,0], green as [0,1,0], and a purple (red + blue) as [1,0,1]. In these examples, we use 0 to mean no light, and 1 to be the maximum light that can be produced by an LED or other colored light emitting device. All colors that a screen can render are within this 3-dimensional space, called the gamut. You might think that is comprehensive, but natural light conditions have a far great range. The reason you are not noticing the difference between seeing a real sunset and one on a screen is that your visual system has a high dynamic range. This means it is adaptive, normalizing inputs to some perceptual range in which the sunset on the screen looks remarkably like a real one, even though in physical reality, they are nothing alike in terms of the electromagnetic energy they emit. But high-dynamic ranging or not, is the brain encoding color similarly to how a monitor is encoding it? The short answer is no. The long answer, which will hopefully lead you to an intuition on manifolds, is as you may have guess, found below.

#im ../assets/figures/019/019-12.png 50 192 Figure 19.12 - The different planar sections of CIELAB space. In the center, the isoluminant plane in which only A and B vary, but L (lightness) is kept constant. Hue varies with angle of a vector originating from the origin (a neutral gray), and the distance from the origin reflect saturation. To the left, the plane created by varying B and L while keeping A constant, to the right, the plane create by varying A and L while keeping B constant.

#h3 19.9 - Representing color in the mind

#pg When human observers are asked to organize a set of colors, we typically find that they will do so based on the three perceptual qualities we described above: lightness, saturation, and hue. Different experimental approaches yield slightly different results and descriptions of the axes along which people seem to sort colors. But they typically are close variants to the axes representing luminance or brightness/lightness, saturation, and hue. Combined, the axes form a color space. RGB is also a color space, but one that is specified in three primary colors screen currently still mix. CIELAB on the other hand is a well-known example of a perceptual color space. It uses quantities L, A and B as a coordinate system, in the same way a monitor uses R, G and B. L stands for lightness, the perceptual quality of luminance. Axes A and B are orthogonal to L and orthogonal to each other. The A axis moves from a turquoise green through a center gray point to a purplish red, and the B axis moves between a pale blue, through a center gray point to a darkish yellow. Why these two axes in particular? This has a lot to do with some additional neural processes that happen after the light coming into your eyes has been captured and processed by your three photoreceptors. If you are truly interested, you can find some references at the end of this chapter that explain this in more detail. The angle on the plane formed by the A and B axes captures changes in hue, the distance from the center of this plane (which is a neutral gray is a measure of saturation. And perpendicular to this plane is the lightness axis. We trust that Figure 19.12 should makes this far clearer than words can ever convey.

#h3 19.10 - The Manifold of hue and saturation in RGB space

#pg Back to geometry and machine learning. The key insight here is that color spaces like RGB and LAB are related through geometric transformation. It is the same kind of transformation we discussed already of how in machine learning, geometrically speaking, inputs are transformed into outputs. And now we are finally at the idea of a manifold. With all the above, we can now ask ourselves how the 2-dimensional color wheel that varies in hue (angular component) and saturation (distance from the center gray point) is embedded in the 3-dimensional space of RGB values a screen produces. That is the core idea of a manifold: a membrane with a dimensionality lower than our original input stimulus space along which some quality of interest varies smoothly and differently defending on the direction one moves along the manifold. As you can see, the manifold of constant lightness but variations in hue and saturation are a complex function of the RGB values. The manifold is both rotated relative to all three axes in RGB space, and curves smoothly.

#im ../assets/figures/019/019-13.gif 50 192 Figure 19.12 - The manifold of hue and saturation as specified by the LAB color space. The three axes represent the intensity of the three primary colors: red, green and blue. The surface with in the cube these three axes create reflects the manifold of hue and saturation, at a fixed lightness or luminance level. This manifold therefore represents a true subspace in 2 dimensions, compared to the 3 dimensions of red, green, and blue.

#h3 19.11 - Relationship with other algorithm objectives

#pg It is this final observation and statement in the previous paragraph that will hopefully convince you of how manifolds are a generalization or superclass of the algorithmic objectives we have already encountered: regression, dimensionality reduction, clustering, and classification. Starting with a gently curved 2-dimensional manifold embedded in 3D color space, we can imagine transforming its lattice in a variety of ways. First, what happens if we imagine the lattice to not be uniform. Imagine again we are asking human observers to sort a set of colors. But this time, we are not asking them to organize them based on perceived proximity but instead on some natural grouping. All the colors the human observer deems to be the 'same' are put in a pile together, and for the moment let us imagine that we instruct the observer to create 8 piles in total.

#h3 19.12 - Faces as an example

#h3 19.13 - Demos

#h3 19.14 - References