CMSC 426

Spring, 2014

Pre-Assessment

There are a number of mathematical concepts that we will make use of in this class. Most of these are things that you learned in math classes such as Calculus or Probability, or earlier in high school, but that you might not remember immediately. I'll review these during the first week of class. After that, I'll assume that you know everything in this document.

1) Derivatives
    a. General properties of the derivative of a function
        i. The derivative tells us how rapidly a function is changing.

$$\frac{df}{dx} = \lim_{\Delta \to 0} \frac{f(x+\Delta) - f(x)}{\Delta}$$

        ii. If a point is a (local) maximum or minimum of a function, the derivative at that point will be zero.
    b. Derivatives of polynomials

For example if: $f(x) = x^3 - 5x + 3$ then: $f'(x) = 3x^2 - 5$

    c. Partial Derivatives

If we have a function of multiple variables, we can take the derivative with respect to just one of them. In doing this, we treat the other variables as constants, and see how the function changes when just this one variable changes. For example, if:

$$f(x, y) = 3xy^2 - 2x^2y^3 - 7xy + 3$$

Then:

$$\frac{\partial f}{\partial y} = 6xy - 6x^2y^2 - 7x$$

2) Integrals

a. Single integrals

A (Riemann) integral is just the limit of dividing a function into little vertical strips and adding up their area. It is the continuous version of summation. Recall that it reverses the effects of a derivative. For example:

$\int 3x^2 - 5dx = x^3 - 5x + K$(where K is any constant)

b. Double integrals

When we have a function of two variables, we can integrate over both of them. When we integrate over one variable, we treat the other variable as if it were a constant. For example,

$$\int\int 6xy^2 - 8xy + 3dxdy = x^2y^3 - 2x^2y^2 + 3xy + K_1x + K_2y + K_3$$

3) Matrices and vectors

A vector is a list of numbers, which we usually interpret as a direction and length. For example, the vector (1,2) can be thought of as an arrow from the origin (0,0) to the point (1,2). Equally, it provides the direction and length from (7,3) to (8,5). We will also talk about column vectors, notated as $(1,2)^T$, or:

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

A vector can also be thought of as a matrix with one row, while a column vector is a matrix of one column. We can scale a vector by multiplying all its elements by the same constant value. s(x,y) = (sx,sy), which has the same direction as (x,y), but is longer by a factor of s. We can add two vectors by adding their elements, so (x,y) + (w,z) = (x+w, y+z).

a. Normalizing a vector

The length of a vector is given by the Pythagorean theorem. Denoting the length of a vector as ||(x,y)|| we have: $\|(x,y)\| = \sqrt{x^2+y^2}$. If the length of a vector is 1, we call it a unit vector. In order to scale a vector so that it has length one, we just divide all its elements by the length. So: $\frac{(x,y)}{\|(x,y)\|}$ is a unit vector. Often, we denote a vector with an arrow over it, and a unit vector with a hat, as: $\vec{x}$ or $\hat{x}$

b. The inner product

The inner product (aka the dot product) is one way to multiply two vectors together and produce a scalar value. We just multiply all the corresponding components of the vectors together and add them up. For example:

$$(x_1, y_1) \cdot (x_2, y_2) = x_1 x_2 + y_1 y_2$$

c. Multiplication of matrices and vector

If A is a matrix, we may refer to the element in row *i* and column *j* as $a_{i,j}$. Then we can multiply two matrices together as:

$$(AB)_{i,j} = \sum_{k=1}^{n} A_{i,k} B_{k,j}$$

where $(AB)_{i,j}$ denotes the element in AB in row *i* and column j. Another way to think about this is that $(AB)_{i,j}$ is the inner product between the *i*'th row of A and the *j'th* column of B. Note that for this to work out, the number of columns in A and the number of rows in B has to be the same. Note that in the same way, we can multiply a matrix by a column vector, which always produces a new column vector. Also note that matrix multiplication is associative, but not commutative, that is, A(BC) = (AB)C, but in general, $AB \neq BA$.

d. Eigenvalues and eigenvectors

A non-zero vector is called an *eigenvector* of a matrix if multiplying the vector by the matrix scales it, but doesn't change its direction. So, for example, *x*, (a column vector) is an eigenvector of the matrix A if and only if $Ax=\lambda x$ for some scalar value λ.

e. Matrix inverse.

The identity matrix, I, is kind of like the number 1. Multiplying a matrix by I doesn't change it. That is, for any matrix A, AI=IA=A. I is a square matrix whose diagonal elements are all 1, and whose off-diagonal elements are all 0.

If we denote the *inverse* of the matrix A as A^{-1} then $AA^{-1}=A^{-1}A=I$.

4) Trigonometric functions and their derivatives

Recall that if we have a right triangle with a hypotenuse of length one, and one of the corners of the triangle has an angle of θ. Then the side opposite the corner has a length sin q and the side adjacent to the corner has a length $\cos\theta$.

It's important to note that if x and y are unit vectors, and the angle between them is θ then the inner product between x and y (which we can also denote <x,y>) will be equal to $\cos\theta$.

5) Linear equations

a. Equation for a line in 2D.

We can express the equation for a line in 2D as a single, linear equation. For example we can write y = mx+b (this doesn't work for vertical lines, which could be written as x = b). This formula has an interpretation of m as the slope of the line and b as the height at which the line intersects the y axis.

We can also describe a line using a starting point and a vector. For example, we can write:

$$(x, y) = (x_0, y_0) + t(u, v)$$

Here we think of x_0, y_0, u, and v as constants that are specific to the line. As t varies, we get different points on the line. We can think of (x_0, y_0) as the start of the line. If (u,v) is a unit vector, then t tells us how far a point is from the starting point. We can also think of this as expressing a line with two equations, one involving x and t, and the second involving y and t.

b. Equation for a line in 3D

In 3D, a single linear equation, like Ax+By+Cz+D=0 (where A,B,C, and D are scalars) describes a plane, not a line. We can describe a line with two linear equations of this form, because the two equations describe two planes, which will intersect in a line.

We can also describe a line in 3D in much the same way that we did in 2D, using an equation:

$$(x, y, z) = (x_0, y_0, z_0) + t(u, v, w) \quad (1)$$

c. Intersection of a line and a plane

Since a plane is described with one linear equation with three unknowns, x, y, and z, and a line in 3D is described with two linear equations, any point that lies on the intersection of a line and a plane must satisfy all three equations. Generally there is a single point that does this, although it is also possible that the three equations have no solution (i.e., the line is parallel to the plane and doesn't intersect it) or many solutions (the line is in the plane, and all points on the line are in their intersection).

If we describe a line using a starting point and vector, as in Equation (1) above, we can think of that as three equations with four unknowns. The plane is described by an equation also, so we wind up with four equations with four unknowns, which generally gives us a point.

6) Histograms

A histogram is a way of counting the number of occurrences of different values of some variable. Suppose, for example, we looked at the temperature every day last year. This would give us 365 numbers. A histogram would tell us how many times the temperature was 0 degrees, 1 degree, etc… during the last year. This gives us a way of characterizing a large amount of data in a compact form.

7) Probability

a. Bayes law

If A is an event, we write the probability of A occurring as P(A). If B is another event, we write P(A,B) for the probability of both A and B occurring, and we write the probability of A occurring if we know that B will occur as P(A|B). Since the probability of both A and B occurring can always be described as the probability of B occurring, and the probability of A occurring conditioned on B's occurrence, we have the identity: P(A,B) = P(A|B)P(B) = P(B|A)P(A). Dividing both sides of the last two expressions by P(B) we have: P(A|B) = P(B|A)P(A)/P(B). This is Bayes law.

b. Variance, standard deviation

If x is a random variable, the expected value of x is written E(x). This is, in some sense, the average behavior of x. For example, if x is a discrete valued random variable that takes on the possible values $x_1, x_2, \ldots, x_n$ then we have:

$$E(x) = \sum_{i=1}^{n} x_i \, P(x = x_i)$$

The variance of x, var(x), measures how much x tends to deviate from its expected value. For example, if x always has the same value, then var(x) = 0. Generally:

$$var(x) = E\left(\left(x - E(x)\right)^2\right)$$

The standard deviation is the square root of the variance.

c. Normal/Gaussian distribution

The Normal distribution gives rise to continuous values that form a bell curve. In vision we always call this a Gaussian distribution. Lots of random variables have a Gaussian distribution because of the law of large numbers, which shows that if a random variable is really the sum of a bunch of independent effects that have almost any distribution, none of them so big as to dominate the sum, then that variable will have a Gaussian distribution. This has the form:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

where σ is the standard deviation of the Gaussian and μ is the mean.