Review Notes: The General Case

Spencer Martz

October 4, 2024

1 Vectors and R^3

1.1 Vector Operations

1.1.1 Dot product, $\vec{a} \cdot \vec{b}$

The dot product of two vectors is a operation which multiplies each pair of terms in the vectors together then sums the result, returning a scalar value. There are a number of ways to interpret the dot product, but it primarily encodes the angle between two vectors.

$$\vec{a} \cdot \vec{b} = a_1b_1 + a_2b_2 \cdots = ||\vec{a}|| \; ||\vec{b}|| \cos\theta, \text{where } \theta \text{ is the angle between a and b.} \tag{1}$$

Two vectors are orthogonal if their dot product is 0.

1.1.2 Cross product, $\vec{a} \times \vec{b}$

The cross product of two vectors is an operation which returns another vector which is normal to the original two. It is calculated by expansion of the following determinant:

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_2 \\ b_1 & b_2 & b_3 \end{vmatrix} \tag{2}$$

The norm of the cross product $||\vec{u} \times \vec{v}||$ is equal to the area of the parrallogram spanned by $\vec{u}$ and $\vec{v}$, and it's half is equal to triangle spanned by the same. Also, the parallelpiped spanned by u, v, and w is given by $|(\vec{u} \cdot (\vec{v} \times \vec{w}))|$, or equivalently $det(u, v, w)$

1.2 Lines, Planes and Quadric Surfaces

The equation of the line is given by $\vec{r}(t) = P + \vec{v}(t)$, where v is the direction of the line, and P is a point on the line. In the plane, any two nonparallel lines will intersect, but this is not the case in r^3 and beyond. We say that two lines (and generally paths) $\vec{r_1}(t)$ and $\vec{r_2}(t)$ intersect if $\vec{r_1}(t) = \vec{r_2}(s)$, for some s which need not equal t. In the case that they are both equal at some time t, they collide.

The simplest surface that exists in R^3 is the plane, and as such it is worth studing in detail. It may be thought of the set of linear combiations of two basis vectors, and is most usefully encoded by it's "normal" vector, or the vector which is perpendicular (everwhere!) to the plane. This vector determines the orientation of the plane in R^3. As such, if we know one point on the plane, say $p =< x_0, y_0, c_0 >$, then we may completely describe the plane with it's normal vector with the equations:

$$< x - x_0, y - y_0, z - z_0 > \cdot < a, b, c >= 0$$
$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$
$$ax + by + cz = d, \text{where } d = ax_0 + by_0 + cz_0$$

There are also a number of other simple surfaces that appear often in R^3, including:

1. ellipsoids $(\frac{x}{a})^2, (\frac{y}{b})^2, (\frac{z}{c})^2 = 1$
2. hyperbolic and elliptic paraboloids, $z = x^2 - y^2$,$z = x^2 + y^2$
3. one-sheet hyperboloids, cones, or two-sheet hyperboloids,
4. parabolic or hyperbolic, elliptic cylinders– we consider any 3D space with equal cross sections as cylindrical.In these equations only two variables are present (since the other may freely vary).

These are named after the apperance of their traces (in the coordinate planes), with the most common conic section taking the primary name of the surface.

2 Vector Calculus

2.1 Integrals and derivatives of VVF

Integrals and derivative of vector valued functions (of a single variable) are evaluated componentwise. For integration, there is still a constant, however it will be a vector and not a scalar.

2.2 Arc length

For some curve C, let r(t) directly traverse the path from. The the length of C from as a function of t is given

$$s(t) = \int_0^t ||\vec{r}\,'(u)||du \tag{3}$$

If we then desire to parametrize the path with a "arc length parametization", we must find the inverse of s(t), then substuite $t = s^{-1}(t)$ into the original parametrization.

2.3 Motion in R^3

As in single variable calculus, integrals and derivatives of vector valued functions have a interpretation of movement. Suppose we are given a vector path $\vec{r}(t)$ which represents an objects path of travel. Upon differentiating (with respect to t), we get $\vec{r'}(t)$, which is called the tangent or velocity vector of the object (at time t).

This vector encodes both the direction and magnitude of travel at time t– it's length, $||\vec{r'}||$ is the speed at which the object travels at time t. Now we may differentiate $\vec{r'}(t)$ to get $\vec{r''}(t)$, the acceleration vector. As in single variable caluclus, this vector encodes the change in velocity at a time T, but there is an additional complexity.

We may also integrate. Given a vector valued acceration, it's integral will be the velocity of the object, upto a constant vector, which is often the inital velocity (since the function will vanish just leaving the constant at t=0). We may integrate again, which gives the r(t) position term as well as another constant vector, which is also often the intial position. In summary, letting $\vec{a}(t)$ be the accelation, we have:

$$\int \vec{a}(t) = \vec{v}(t) + \vec{v_0}, \int \vec{v}(t) = \vec{r}(t) + \vec{r}_0 \tag{4}$$

2.4 Tangent and Normal acceleration decomposition

In single variable calculus, it is an (occosianally) useful fact that the derivative of a constant function is 0. And this idea logical extends to the derivative of velocity: a constant velocity has zero accelation. And in some sense, that is still true for our purposes. There is no change in speed for a constant... speed, but their may be a change velocity and the acceleration vector. This is because we may change direction while still maintaining our speed, which will involve changing our velocity vector, but not it's magnitude. As such it is useful to decompose the accelation vector, into its "tangent" and "normal" components, which represent the change in speed and change in direction, respectively.

Suppose as before we have a vector valued function representing the path of the object, and let $\vec{r}(t)$, $\vec{v}(t) = \vec{r}\,'(t)$, $\vec{a} = \vec{r}\,''(t)$ denote the position, velocity, and accelation vectors for the object at a time t. Addionally, let the scalar $v(t) = ||\vec{v}(t)||$ be the speed of the object. We wish to decompose $\vec{a}(t)$ into two components:

$$\vec{a}(t) = a_T\vec{T}(t) + a_N\vec{N}(t) \tag{5}$$

where a_T and a_N are the (scalar) "tangential" and "normal" components of accelation, and $\vec{T}$ and $\vec{N}$ are unit tangent and normal vectors. It is best to understand this first through orthogonal projection. We desire $a_T\vec{T}$ to be the "part" of accelation in the same direction as the velocity (the direction of travel), and $a_N\vec{N}$ to the the part of accelation is perpendicular to the direction of travel. It follows that the (former) vector is equal to the orthogonal projection of $\vec{a}$ onto $\vec{v}$, expressed with dot products as

$$a_T\vec{T} = proj_{\vec{v}}(\vec{A}) = \frac{\vec{a}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v},\text{ and so } a_N\vec{N} = \vec{a} - proj_{\vec{v}}(\vec{A}) \tag{6}$$

Now, noting that $\vec{v}\cdot\vec{v} = ||\vec{v}||^2$, we expand the projection as:

$$\frac{\vec{a}\cdot\vec{v}}{||\vec{v}||}\frac{\vec{v}}{||\vec{v}||} \tag{7}$$

The second term is a unit vector in the direction of V, so we take

$$T\vec{(t)} = \frac{\vec{v}(t)}{||\vec{v}(t)||} = \frac{\vec{v}(t)}{v(t)},\text{ so } a_T = \frac{\vec{a}\cdot\vec{v}}{||\vec{v}||} = \vec{a}\cdot\vec{T} \tag{8}$$

3 Multivariable functions

3.1 Limits

Limits are similar in multiple variables, but there is an important discinction. For $f(x, y)$ (say), the limit of f as (x,y) approaches (a,b) exists iff all paths approach that limit. This is analougous to the two sided defintion of the limit existing in multiple variables-as we come in from any direction (or any path, more precisely), the function should approach the same value. (Example?)

3.2 Partial Derivatives and the Gradient

For some function f(x,y), the partial derivative of f with respect to x, denoted f_y or $\frac{\partial f}{\partial y}$ is defined as:

$$\frac{\partial f(x,y)}{\partial x} = \lim_{h \to 0} \frac{f(x+h,y) - f(x,y)}{h} \text{ and similarly,} \tag{9}$$

$$\frac{\partial f(x,y)}{\partial y} = \lim_{h \to 0} \frac{f(x,y+h) - f(x,y)}{h} \tag{10}$$

This represent the change in f with respect to x, and all other variables constant. This can be interpreted as the slope of a tangent vector on the surface of f=z, paticuarly with f_x coplaner to the xz axis, and f_y coplanar to the xy axis as in the figure. As such, for a point (a, b, f(a,b)), we may take $< 1, 0, f_x(a, b) >$ and $< 0, 1, f_y(a, b) >$ to develop a tangent plane equation, with normal

$$\vec{n} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 0 & f_x(a,b) \\ 0 & 1 & f_y(a,b) \end{vmatrix} =< f_x(a,b), f_y(a,b), -1 > . \tag{11}$$

Thus our plane equation is

$$< x-a, y-b, z-f(a,b) > \cdot < f_x(a,b), f_y(a,b), -1 >= 0 \tag{12}$$

$$0 = f_x(a,b)(x-a) + f_y(a,b)(y-b) - z + f(a,b) \tag{13}$$

$$z = L(x,y) = f_x(a,b)(x-a) + f_y(a,b)(y-b) + f(a,b) \tag{14}$$

This $L(x, y)$ is also known as the linearization of f at (a,b), and is approximatiely equal to f near f(a,b). So to approimate some small change in f with respect to small changes in x and y (from (a,b)), denoted $\Delta x, \Delta y$, we take

$$L(a+\Delta x, b+\Delta y) = f_x(a,b)(a+\Delta x - a) + f_y(a,b)(b+\Delta y - b) + f(a,b) \tag{15}$$

$$= f_x(a,b)\Delta x + f_y(a,b)\Delta y + f(a,b) \tag{16}$$

Now if we denote the change in f by $\Delta f(x, y) = f(x, y) - f(a, b)$, then the linear approximation forumla follows:

$$\Delta f(\Delta x, \Delta y) \approx f_x(a,b)\Delta x + f_y(a,b)\Delta y \tag{17}$$

3.3 Direction derivatives and the Tangent Plane

The directional derivative is a natural extension of the partial derivative-the represent the slopes of tangent lines in other directions then x and y. However they are more easily understood through a less natural idea, that of the gradient. The gradient of a function is the the vector of it's partial deriviatives– say, for f(x,y),

$$\nabla f =< \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} >=< f_x, f_y > \tag{18}$$

There is an important theorem related to the gradient, called the "chain rule for paths". It states that for a path r(t), the derivtiave of f(r(t)) with respect to t is given:

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t) \tag{19}$$

Now we are ready to introduce the directional derivative. Let $\vec{u} =< h, k >$ be a (unit!) vector. Thus the path $r(\vec{t}) =< a + ht, b + kt >$ is a line in the direction of $\vec{u}$. The defintion of the directional derivative is then:

$$D_{\vec{u}} = \frac{\mathrm{d}}{\mathrm{d}t} f(\vec{r}(t))\big|_{t=0} = \lim_{t\to 0} \frac{f(a + ht, b + kt) - f(a, b)}{t} \tag{20}$$

Thus by the chain rule for paths, for P=(a,b)

$$D_{\vec{u}} = \nabla f(P) \cdot \vec{r}\,'(t) = \nabla f_P \cdot \vec{u} \tag{21}$$

This relation allows us to form an interpretation of the gradient. Suppose ∇f is non-zero,and let θ be the angle between $\vec{u}$ and ∇f, then

$$D_{\vec{u}} = \nabla f_P \cdot \vec{u} = ||\nabla f_P||||\vec{u}|| \cos\theta \tag{22}$$

Noting $\vec{u}$ is a unit vector, and addtionally that cosine is bounded by -1 and 1, we see:

$$-||\nabla f_P|| \le D_{\vec{u}} \le ||\nabla f_P|| \tag{23}$$

In other words, the maxmium possible rate of change at P is given by $||\nabla f_P||$. Additionally, this rate of changed is achieved when $\cos\theta = 1$ – when the angle between u and ∇f is 0, or when u is in the direction of ∇f,

3.4 General chain rule and implict differentiation

The general chain rule is a method to more easily calculate the deriviavte of composed functions. Say we have a function $f(x_1, x_2 \dots x_n)$ of n variables, where each x_i is a differentiable function of m independent variables $t_1, t_2, \dots t_m$. Then the chain rule states that

$$\frac{\partial f}{\partial t_k} = \frac{\partial 1}{\partial x_1}\frac{\partial x_1}{\partial t_k} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_k} + \dots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_k} = \nabla f \cdot < \frac{\partial x_1}{\partial t_k}, \dots, \frac{\partial x_n}{\partial t_k} > \tag{24}$$

Now suppose we have an implicit function F(x,y,z)=0– a surface, in R^3. Then$z = z(x, y)$; it is a function of x and y.We may not be able to express z(x,y) easily, but $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ can be found. First we use the chain rule to differentiate F (wrt for example), treating z as a function of x and y.

$$\frac{\partial F}{\partial x} = \frac{\partial F}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0 \tag{25}$$

Noting $\frac{\partial x}{\partial x} = 1$ and $\frac{\partial y}{\partial x} = 0$ (y does not depend on x), we have:

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = F_x + F_z\frac{\partial z}{\partial x} = 0 \tag{26}$$

Thus (the $\frac{\partial z}{\partial y}$ case is similar)

$$\frac{\partial z}{\partial x} = \frac{-F_x}{F_z}, \quad \frac{\partial z}{\partial y} = \frac{-F_y}{F_z} \tag{27}$$

3.5 Optimization

Optimization in multiple variables is not unlike in single variables. The fundmental idea that extrema occur either at endpoints, or where the derivative is equal to zero is still true, but there is a new test to determine the nature of the extrema.

3.5.1 Local extrema

Local extrema occur at critical points, where f_x or f_y equal 0. These points may be local minmum, maximums, saddle points, or none of these. The Hessian determinant is a test (much like the 2nd derivative test in SVC) which determines the nature of these points. It is given by:

$$D(a,b) = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx}(a,b)f_{yy}(a,b) - f_{xy}^2(a,b) \tag{28}$$

D	f_{xx}	Point type
$D > 0$	$f_{xx} > 0$	Local minimum
$D > 0$	$f_{xx} < 0$	Local maximum
$D < 0$	doesn't matter	Saddle point
$D = 0$	doesn't matter	indeterminate

3.5.2 Global extrema

Often we are more interested in the maximum a function takes on, not just the local extreme values. It is important to note for some, or even many functions, there is no extrema an open interval. However if we choose our domain to be closed and bounded, the following is garunteed.

Theorem: Extrema Let f(x,y) be continous on a closed and bounded domain $D \subset R$. Then:

1. f(x,y) takes on a minimum and maximum value on D
2. The extrema occur either at the critical points of f, or along the boundary of D.

The method for finding global extrema follows: First, find and evaluate and crtical points within the boundary of D. Second, parametrize the boundary of D with a function r(t), and use SVC to find the extrema on the endpoints. Finally, compare and determine the global minimums and maxima.

3.6 Lagrangian Optimization

Lagrangian optimazation is a method of solving optimization problems given a contraint (one example of this occurs when solving for the boundary in a global extrema problem). For a constraint function g(x,y)=0, we say the extrema of f(x,y) occurs at points where the Lagrange condition holds

$$\nabla f = \lambda \nabla g, \text{ or } \begin{matrix} f_y(a,b) = \lambda g_y(a,b) \\ f_x(a,b) = \lambda g_x(a,b) \end{matrix} \tag{29}$$

From this we obtain a system of 2 equations in 3 variables– we solve for lambda, and obtain a relation for x and y. This is then substitued into the constraint equation to produce the (x,y) points which satify all three equations–that is, they are on the contraint, and satify the lagrange condition (and are thus minimums or maxmimums). In essence, this method allows optimization problems to be solve more simply with algebra.