Vector Analysis

Onkar Pandit

04 March 2022

Last updated on June 4, 2022.

Contents

1 Introduction

2 Basics

A vector is a mathematical object which denotes magnitude and direction (e.g. velocity, force etc.) whereas scalar has only magnitude no direction (e.g. mass, time, etc.).

Scalar fields If for each point x, y, z in space R, a scalar number $\phi(x, y, z)$ is associated then ϕ is called scalar function of position and ϕ is called scalar field defined on space R. **does the scalar function has to produce unique scalars?** A scalar field which is independent of time is called stationary or steady-state scalar field.

For example, temperate in 3 dimensional space (a temperature field) can be written as $T(x, y, z)$ – a scalar function of the position. An example of a scalar field in electromagnetism is the electric potential.

Vector field If for each point x, y, z in space R, a vector corresponds $v(x, y, z)$, then v is called vector function of position and we say that vector field v is defined over space R. A vector field which is independent of time is called stationary or steady-state vector field.

For example, $v(x, y, z) = x^2yi - yzj + xz^2k$ is a vector field. Another vector field $v(x, y) = xi + yj$ is called *source field* as vector are going out of the origin on the opposite of that $v(x, y) = -xi - yj$ is called *sink field* as vectors are pointed towards origin.

About the vector field, we can also say, it is a function which takes same number of inputs and produces equal number of outputs. Thus, at each point in the space it gives the direction and magnitude of the field as shown in Figure 1.

3 Vector-valued Functions

A vector-valued function is a function of the form

$$r(t) = f(t)i + g(t)j \qquad \text{or } r(t) = f(t)i + g(t)j + h(t)k$$

where f, g, h are real-valued functions of parameter t and i, j, k are unit vectors. Vector-valued functions are also written in the form $r(t) = \langle f(t), g(t)\rangle$ or $r(t) = \langle f(t), g(t), h(t)\rangle$. It is a function from the real numbers R to the set of all two- or three-dimensional vectors.

Graphing Vector-Valued Functions A plane vector consists of two quantities: direction and magnitude. Given any point in the plane (the initial point), if we move in a specific direction for a specific distance, we arrive at a second

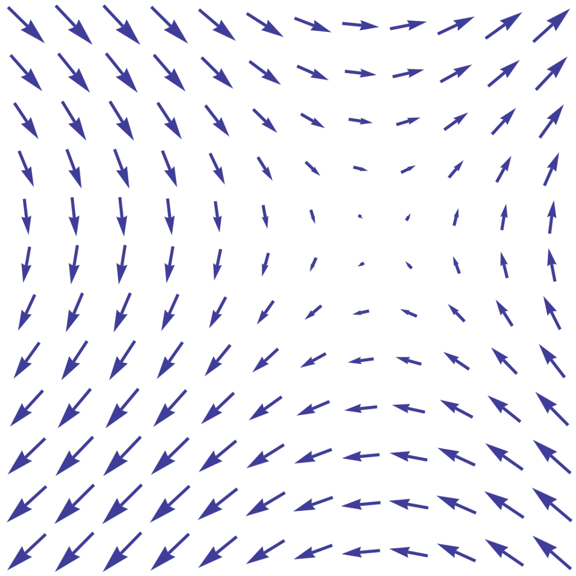

Figure 1: At each point in the two-dimensional plane the vector field gives direction as well as magnitude of the field.

point. This represents the terminal point of the vector. We calculate the components of the vector by subtracting the coordinates of the initial point from the coordinates of the terminal point.

A vector is considered to be in standard position if the initial point is located at the origin. When graphing a vector-valued function, we typically graph the vectors in the domain of the function in standard position, because doing so guarantees the uniqueness of the graph. This convention applies to the graphs of three-dimensional vector-valued functions as well.

The graph of a vector-valued function of the form: $r(t) = f(t)i + g(t)j$ consists of the set of all points $\langle f(t), g(t) \rangle$, and the path it traces is called a **plane curve**.

The graph of a vector-valued function of the form: $r(t) = f(t)i+g(t)j+h(t)k$ consists of the set of all points $\langle f(t), g(t), h(t) \rangle$, and the path it traces is called a **space curve**.

Any representation of a plane curve or space curve using a vector-valued function is called a **vector parametrization** of the curve. Each plane curve and space curve has an orientation, indicated by arrows drawn in on the curve, that shows the direction of motion along the curve as the value of the parameter t increases.

Let us see examples of the vector-valued functions and their graphs.
Graph for vector-valued function: $r(t) = 4\cos ti + 3\sin tj \quad 0 \leq t \geq 2\pi$ is shown in Figure 2.
Graph for vector-valued function: $r(t) = 4\cos ti + 3\sin tj + tk \quad 0 \leq t \geq 2\pi$ is

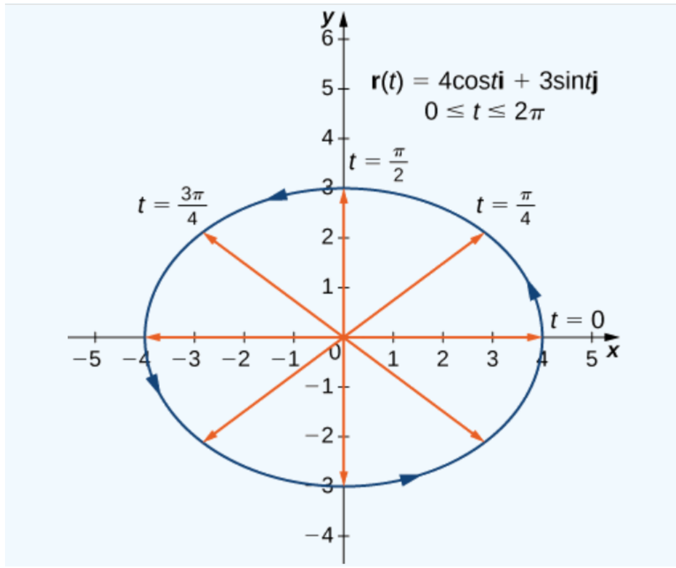

Figure 2: Graph of $r(t) = 4\cos ti + 3\sin tj \quad 0 \leq t \geq 2\pi$.

shown in Figure 3.

There is a similarity between vector-valued functions and parametrized curves. Indeed, given a vector-valued function $r(t) = f(t)i+g(t)j$ we can define $x = f(t)$ and $y = g(t)$. The graph of the parametrized function would then agree with the graph of the vector-valued function, except that the vector-valued graph would represent vectors rather than points.

4 Vector Differentiation and Integration

4.1 Differentiation of vectors

Let $R(u)$ be a vector depending on the scalar function u, then the differentiation of vector with respect to u is given as:

$$\frac{d}{du}R(u) = \lim_{\Delta u \to 0} \frac{\Delta R}{\Delta u} = \lim_{\Delta u \to 0} \frac{R(u + \Delta u) - R(u)}{\Delta u} \tag{1}$$

The differentiation of vector can again produce vector, so it is possible to take its differentiation again. This way, we can find higher order derivatives as $\frac{d^2}{du^2}R(u)$, etc.

Suppose, vector $r(u) = x(u)i + y(u)j + z(u)k$ gives position depending on the values of x, y, z then its derivative w.r.t. u can be given as:

$$\frac{d}{du}R(u) = \frac{dx}{du}i + \frac{dy}{du}j + \frac{dz}{du}k \tag{2}$$

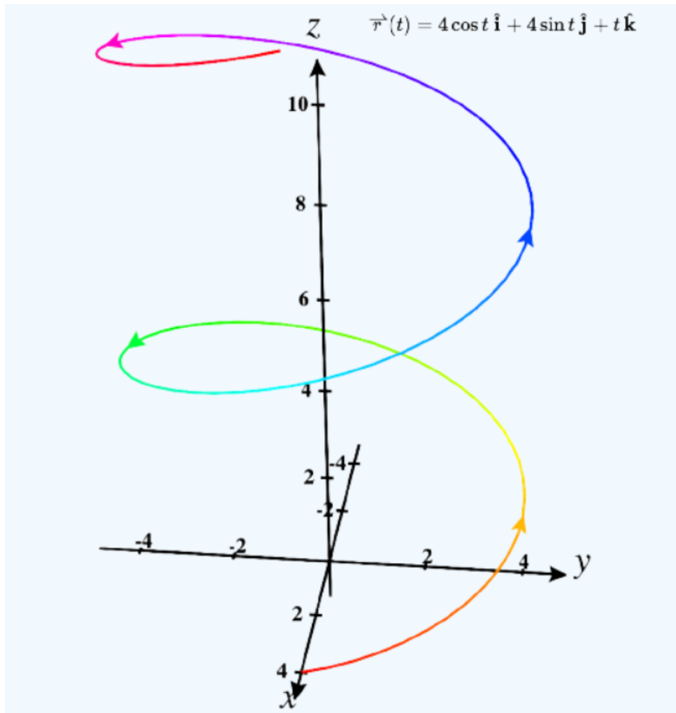

Figure 3: Graph of $r(t) = 4\cos ti + 3\sin tj + tk \quad 0 \leq t \geq 2\pi$. This is also called as *helix*. The k term is always increasing otherwise it is a circle with radius 4.

The above formula looks intuitive and simple as we have seen in the single valued functions. But, there is a little more to it than just addition of the changes. The formula says that the total change in vector is addition of change in x direction, y direction, and z direction. This is the interpretation which is important to be comfortable with.

Continuity and differentiability A vector function $R(u) = R_1(u)i + R_2(u)j + R_3(u)k$ is said to be continuous if all the three scalar functions R_1, R_2, R_3 are said to be continuous in u or if $\lim_{\Delta u \to 0} R(u + \Delta u) = R(u)$. Equivalently if $|R(u + \Delta u) - R(u)| < \epsilon$ for $|\Delta u| < \delta$.

The different formulas which are obtained for real-valued functions like multiplication of functions can also be extended to vector-valued functions, so if we take dot product or cross product of vector-valued functions, similar rules as of real-valued functions are applied to get the derivative value.

It is crucial to understand that, vector can contain only single variable encoded in them, for example, when we say a vector $r(u)$ here, we are saying that vector may be having multiple elements, but they all are depending on u, for example, they may be $[u^2 \;\; 2u \;\; 4u^3]$,

which contains three elements but all are depending on u. There can be more involved function $R(u) = (x(u, y(u), z(u))$ which means that a vector depends on x, y, z and they in turn depend on u. Thus, it is important to measure effect of change of value of x, y, z independently to understand the effect on $R(u)$.

Tangent Vectors and Unit Tangent Vectors The derivative at a point can be interpreted as the slope of the tangent line to the graph at that point. In the case of a vector-valued function, the derivative provides a tangent vector to the curve represented by the function.

Consider the vector-valued function $r(t) = \cos ti + \sin tj$, its derivative is given as: $r'(t) = -\sin ti + \cos tj$. If we substitute the value $t = \pi/6$ into both functions we get, $r(\pi/6) = \frac{\sqrt{3}}{2}i + \frac{1}{2}j$, and $r'(\pi/6) = \frac{\sqrt{-}+\frac{1}{2}i+3}{2}j$. The graph of this is shown in Figure 4.

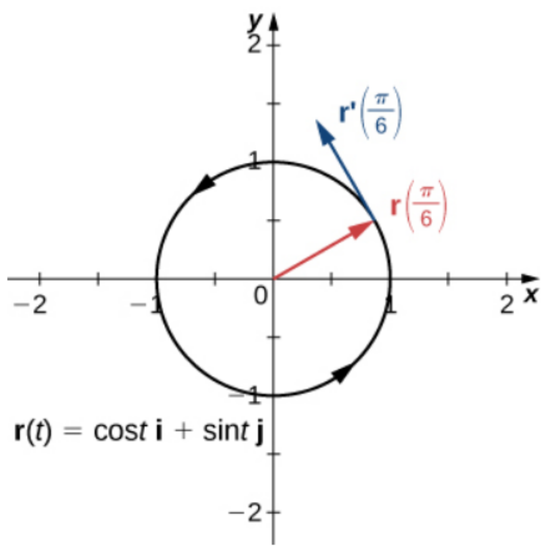

Figure 4: Graph and its tangent line obtained with derivative at that point.

Unit tangent vector is exactly what it sounds like: a unit vector that is tangent to the curve. To calculate a unit tangent vector, first find the derivative $r'(t)$. Second, calculate the magnitude of the derivative. The third step is to divide the derivative by its magnitude.

$$u_r(t) = \frac{r'(t)}{||r'(t)||} \tag{3}$$

Partial Derivatives of vectors Let A be a vector function of x, y, z then partial derivatives w.r.t. these variables is given as:

$$\frac{\partial A}{\partial x} = \lim_{\Delta x \to 0} \frac{A(x+\Delta x, y, z) - A(x, y, z)}{\Delta x}$$
$$\frac{\partial A}{\partial y} = \lim_{\Delta y \to 0} \frac{A(x, y+\Delta y, z) - A(x, y, z)}{\Delta y}$$
$$\frac{\partial A}{\partial z} = \lim_{\Delta z \to 0} \frac{A(x, y, z+\Delta z) - A(x, y, z)}{\Delta z}$$

Similar to function of single variable, the remarks on continuity and differentiability can be extended to function with multiple variables.

Differentials of vector for a function $A(x, y, z)$ can be written as:

$$dA = \frac{\partial A}{\partial x}dx + \frac{\partial A}{\partial y}dy + \frac{\partial A}{\partial z}dz \tag{4}$$

If we want to find out a component of vector in certain other vectors direction, then we take dot product.

4.2 Integration of vectors

The concepts of definite and indefinite integrals which were proposed for real-valued functions, can also be extended to vector-valued functions. Also, just as we can calculate the derivative of a vector-valued function by differentiating the component functions separately, we can calculate the antiderivative in the same manner. Furthermore, the Fundamental Theorem of Calculus applies to vector-valued functions as well.

The antiderivative of a vector-valued function appears in applications. For example, if a vector-valued function represents the velocity of an object at time t, then its antiderivative represents position. Or, if the function represents the acceleration of the object at a given time, then the antiderivative represents its velocity.

Definite and Indefinite Integrals of Vector-Valued Functions Let f, g, h be integrable real-valued functions over the closed interval $[a, b]$.

1. The indefinite integral of a vector-valued function $r(t) = f(t)i + g(t)j$ is

$$\int r(t)dt = (\int f(t)dt)i + (\int g(t)dt)j \tag{5}$$

   Definite integral can be given as:

$$\int_a^b r(t)dt = (\int_a^b f(t)dt)i + (\int_a^b g(t)dt)j \tag{6}$$

2. Similarly, the indefinite integral of a vector-valued function $r(t) = f(t)i + g(t)j + h(t)k$ is

$$\int r(t)dt = (\int f(t)dt)i + (\int g(t)dt)j + (\int h(t)dt)k \tag{7}$$

Definite integral can be given as:

$$\int_a^b r(t)dt = (\int_a^b f(t)dt)i + (\int_a^b g(t)dt)j + (\int_a^b h(t)dt)k \tag{8}$$

5 Gradient, Divergence, Curl, and Laplacian

Gradient For a real-valued function $f(x, y, z)$ on R^3, the gradient $\nabla f(x, y, z)$ is a vector-valued function on R^3, that is, its value at a point (x, y, z) is the vector

$$\nabla f(x, y, z) = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}) = \frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k$$

where each of the partial derivatives is evaluated at the point (x, y, z). So in this way, we can think of the symbol ∇ as being "applied" to a real-valued function f to produce a vector ∇f.

It turns out that the divergence and curl can also be expressed in terms of the symbol ∇. This is done by thinking of ∇ as a vector in R^3, namely $\nabla = \frac{\partial}{\partial x}i + \frac{\partial}{\partial y}j + \frac{\partial}{\partial z}k$, which can be applied over any function. Here, the symbols $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ are to be thought of as "partial derivative operators" that will get applied to a real-valued function, say $f(x, y, z)$, to produce the partial derivatives $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$. For instance, $\frac{\partial}{\partial x}$ "applied" to $f(x, y, z)$ produces $\frac{\partial f}{\partial x}$. Is ∇ really a vector? Strictly speaking, no, since $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ are not actual numbers. But it helps to think of ∇ as a vector, especially with the divergence and curl, as we will soon see.

Directional Gradient It is a gradient or derivative considered in the direction of a given vector. The value of the directional derivative changes as the vector value changes.

$$\nabla_w f(x, y) = w.\nabla f \qquad = w_1\frac{\partial f}{\partial x} + w_2\frac{\partial f}{\partial y} \qquad if w = [w_1 \;\; w_2]^T \tag{9}$$

Divergence Now, we are considering ∇ as a vector, if we take dot product of ∇ with vector or vector-valued function, we can get divergence. Let $f(x, y, z) = f_1(x, y, z)i+f_2(x, y, z)j+f_3(x, y, z)k$ be a vector field or a vector-valued function which produces vectors, then divergence over f is denoted as div f or $\nabla.f$ and it is obtained as

$$
\begin{aligned}
\nabla.f &= (\frac{\partial}{\partial x}i + \frac{\partial}{\partial y}j + \frac{\partial}{\partial z}k).(f_1(x,y,z)i + f_2(x,y,z)j + f_3(x,y,z)k) \\
&= \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} \\
&= \text{div } f
\end{aligned}
$$

It is interesting to note that, divergence produces "scalar" value from a given vector as it can be seen as dot product of two vectors whereas gradient can be seen as opposite operation, which converts "scalar" value into vectors. In fact, gradient can be considered as vector-valued function, which takes in scalar values and produces vectors.

It is beneficial to think divergence with vector fields. Divergence gives idea about the flow from any point in the vector fields. The positive divergence indicates that the net flow from any point in the vector field is outwards, if the divergence is negative then it means the flow is inwards. The zero divergence means nothing happens at the point i.e. net inflow is equal to net outflow.

Curl It is a cross-product of ∇ and vector-valued function $f(x,y,z) = P(x,y,z)i + Q(x,y,z)j + R(x,y,z)k$ which is denoted as $\nabla \times f$ and is given as

$$
\begin{aligned}
\nabla \times f &= \begin{vmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P(x,y,z) & Q(x,y,z) & R(x,y,z) \end{vmatrix} \\
&= (\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z})i - (\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z})j + (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})k \\
&= \text{curl } f
\end{aligned}
$$

Intuitively, curl gives the direction of the rotation of a point in the vector field. Suppose the rotation is counter-clockwise the curl is positive whereas for the clockwise rotation the curl is negative.

Laplacian For a real-valued function $f(x,y,z)$, the gradient $\frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k$ is a vector field, so we can take its divergence over it:

$$
\begin{aligned}
\text{div } \nabla f &= \nabla.\nabla f \\
&= (\frac{\partial}{\partial x}i + \frac{\partial}{\partial y}j + \frac{\partial}{\partial z}k).(\frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k) \\
&= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}
\end{aligned}
$$

Note that this is a real-valued function, to which we will give a special name: **Laplacian**.

For a real-valued function $f(x, y, z)$, the Laplacian of f, denoted by Δf, is given by

$$\Delta f(x, y, z) = \nabla . \nabla f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \tag{10}$$

Often the notation $\nabla^2 f$ is used for the Laplacian instead of Δf, using the convention $\nabla^2 = \nabla . \nabla$.

Laplacian is the divergence of gradient of the function. Geometrically, it is equivalent to second order derivative of the function. The gradient gives the direction of the steepest ascent and the divergence is positive when we have to go farther from the point which means the point is minima. In the case of maxima, in the neighborhood, every gradient will be away towards this point meaning the divergence near the point will be negative. Thus, similar to the second order derivative of the function, laplacian will be negative near the maxima and positive near the minima.

Hessian Suppose we take outer product of ∇f and ∇ we get a different mathametical entity known as Hessian matrix:

$$\nabla \otimes \nabla f = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix}$$

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \, \partial y} & \frac{\partial^2 f}{\partial x \, \partial z} \\ \frac{\partial^2 f}{\partial x \, \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \, \partial z} \\ \frac{\partial^2 f}{\partial x \, \partial z} & \frac{\partial^3 f}{\partial y^2 \, \partial z} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix}$$

Now, we can see a relation between Hessian matrix and Laplacian, the trace of the Hessian matrix is a Laplacian.

Gradient of Vector-valued function So far, we have seen functions of the form $f : R^n \to R$ of n-variables. The gradient of this functions is given as:

$$\nabla_x f = \text{grad } f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix} \in R^{1 \times n} \tag{11}$$

Note that, gradient can be a column vector or row vector, the row or column form is selected based on the ease of calculations.

Example, find gradient of $f(x) = x^T x$.

$$\begin{aligned} f(x) &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= x_1^2 + x_2^2 + \cdots + x_n^2 \end{aligned}$$

$$
\begin{aligned}
\nabla_x f &= \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix} \\
&= \begin{bmatrix} 2x_1 & 2x_2 & \cdots & 2x_n \end{bmatrix} \\
\nabla_x f &= 2x
\end{aligned}
$$

Now, suppose we have a vector-valued function instead of a scalar function which is of the form: $f : R^n \to R^m$. A function accepts $x \in R^n$ and produces vector:

$$
f(x) = \begin{bmatrix} f_1(x_1, x_2, \cdots, x_n) \\ f_2(x_1, x_2, \cdots, x_n) \\ \vdots \\ f_m(x_1, x_2, \cdots, x_n) \end{bmatrix} \in R^m
$$

The collection of first-order partial derivatives of a vector-valued function f is called Jacobian. The Jacobian is $m \times n$ matrix which is defined and arranged as follows:

$$
\begin{aligned}
\mathbf{J} = \nabla_x f &= \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}
\end{aligned}
$$

This is a matrix which means it is a linear transformation. It will transform a point in a space to other vector. This transformation is nothing but a locally linear transformation at a certain point. Though the actual transformation is non-linear. The Jacobian will produce transformations in the locally linear (zoomed in) are. The determinant gives the magnitude with which the times around the points will be squished or stretched. Suppose a function $f_(x, y)$ and $f_2(x, y)$ are non-linear then the transformation will not produce nice parallel lines as the lines will be curvy because of non-linearity.

Example, consider a function $f(x) = Ax$, find the Jacobian of the function.

$$
\begin{aligned}
f(x) &= Ax \\
&= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
J_x(f) &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \\
&= A
\end{aligned}
$$

6 Maxima and Minima in Multivariate function

In single variable calculus maxima and minima are obtained by differentiating the function and equating it to 0. Similarly, at maxima and minima points in case of multi variable function, all the partial derivatives are 0, which means gradient is 0.

7 Lagrange Multipliers

In constrained optimization problems, we maximize (or minimize) an objective function with certain conditions on the variables that govern the optimization function. Solving constrained optimization functions is not easy so they are first converted into unconstrained optimization problems. Lagrange multipliers provide that neat way to do so.

Suppose without a loss of generality, $f(x, y)$ is an objective function and $g(x, y)$ are constraints where x, y are variables whose values we want to find out. Then the values of x, y which maximize the function lie on the surface where both $f(x, y)$ and $g(x, y)$ touch each other. As shown in Figure 5, the contour lines of both the functions meet. *We know, from the theory that the gradients of the functions are perpendicular to contour lines.* Thus, for both the functions, the gradient will be on the same line, which means both gradients will just differ in magnitude and they will be multiples of some constant. We call this constant as Lagrange Multiplier and write it as

$$\nabla f = \lambda \nabla g \tag{12}$$

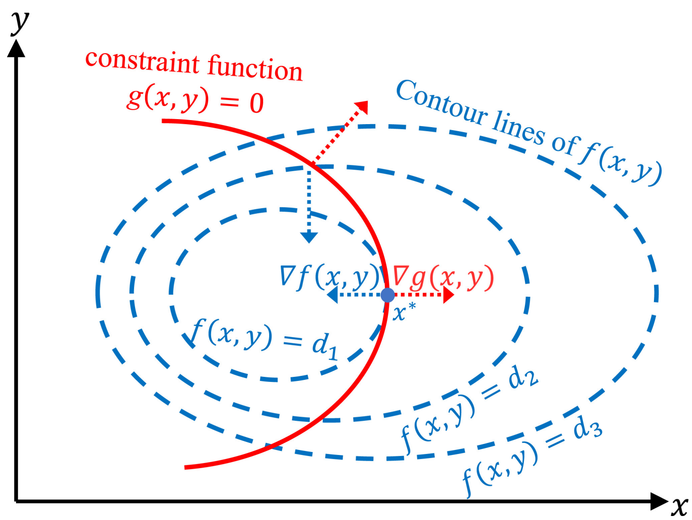

Figure 5: Lagrange Multiplier.

To incorporate these conditions, we create an auxiliary function

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - b)$$

where we introduce b as a new constant for the inequalities so that $g(x, y) - b = 0$ so that the constraints are satisfied. This can be done without a loss generality as any constraints can be converted into such equation.

Now, we maximize this function w.r.t. x, y, λ which means $\nabla_{x,y,\lambda}\mathcal{L} = 0$ which can produce following nice intuition:

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0 \qquad \text{(follows from Eq. 12)}$$
$$\frac{\partial \mathcal{L}}{\partial y} = \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0 \qquad \text{(follows from Eq. 12)}$$
$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 - (g(x, y) - b) = 0 \implies g(x, y) = b \qquad \text{(all the constrains are satified)}$$

Thus, solving $\nabla_{x,y,\lambda}\mathcal{L} = 0$ is going to produce right results given that we maintain $\lambda > 0$ because failing in this condition can violate the Eq. 12.

Lagrange multiplier also has a nice interpretation, it is rate of change in f with respect to change in constraints b:

$$\frac{\partial f}{\partial b} = \lambda$$

. This kind of intuition can be helpful in understanding how constraint can be changed to increase our objective.

Discussion

7.1 Vector vs multi-variable functions

There is a difference between vector-valued functions and multi-variable functions. Suppose, we have a function which we get from two variables and vector-valued function of the form $\langle f(t), g(t) \rangle$, as the vector-valued function has both magnitude as well as direction. The vectors or multi-variable functions which we discuss in machine learning are not vector-valued functions but are simple functions which does not have direction. There is no notion of direction but only magnitude. I guess, that is the subtle difference between vector-valued function and multi variable function.

Another point is that, vector valued functions can contain multi-variable functions as we see that each component of a vector is a multi-variable function in x, y, z. Thus, they are related in this way, but it is not necessary that vector-valued functions use multi-variable functions, they can contain a function of only single variable as well.

7.2 Vector division

We want to think on the point that, can two vectors be divided? This in turn gives rise to the concept of inverse of vectors. These are relatively undefined concepts so there is no direct notion of divisibility of two vectors but we can define simple division. Before, that we have define which kind of multiplication we want to consider in the division operation. Suppose, we want to divide vector u by v i.e. we want a vector value for $\frac{u}{v}$, here assumption is that the division is going to produce vector and not a scalar. Now, suppose, $w = \frac{u}{v}$, which means $u = wv$. This shows, we want to first define, what kind of multiplication, we want to consider. Suppose, we consider a simplest multiplication operation, Hadamard operation, then finding w becomes quite easy, we just divide each element of u by corresponding element in v.