Domain based applications of Multi-variable Calculus

Name:	Debopriya Bose
Registration No.:	2248034
Department:	Department of Statistics and Data Science
University:	Christ(Deemed to be University), Bangalore

Class:	2 M.Sc. Data Science
Semester:	2
Subject:	MDS231
Assignment:	CIA1
Date:	28/01/2022
Professor:	Dr. Ashoka K

Contents

1 Introduction

Calculus with functions of many variables, commonly referred to as multivariate calculus, is an extension of calculus in one variable that involves the differentiation and integration of functions involving multiple variables as opposed to simply one. One way to think of multi-variable calculus is as an introduction to more complex calculus. The term "vector calculus" refers to the specific situation of calculus in three dimensional space.

Multi-variable calculus is one of the fundamental techniques used in Applied Mathematics. It is utilised in a variety of disciplines, including computer graphics, engineering, physical science, and economics. Following are a few examples of multi-variable calculus applications:

1.For dynamic systems, multi-variable calculus offers a tool.

2. It is applied for optimum control in a continuous-time dynamic system. Deriving the formulas to determine the link among the collection of empirical data is helpful while performing regression analysis.

3. High dimensional systems that demonstrate the deterministic nature can be studied and modelled to aid in the study of engineering and social science.

4. To forecast future stock market patterns, quantitative analysts in finance apply multi-variable calculus.

In this report, we are going to broadly discuss some of the applications of multi-variable calculus in Computer Science, Finance and Electrical Engineering domains.

2 Domains

2.1 Computer Science

In order to discover the optimum parameters that provide the lowest prediction error, many of the widely used machine learning techniques employ some variation of gradient descent for optimization, which is essentially an application of multi-variable calculus. The gradient is a multi-variable calculus idea that is utilised in machine learning. The gradient, or direction of the steepest ascent for the surface, is calculated as the partial derivatives of the surface with respect to x and y. Gradient descent is a machine learning technique that seeks to identify the function that best classifies a given set of data. Here, the goal is to reduce the difference in accuracy between the function's output and the real data. In order to do this, a graph of error can be created (also known as cost in machine learning).

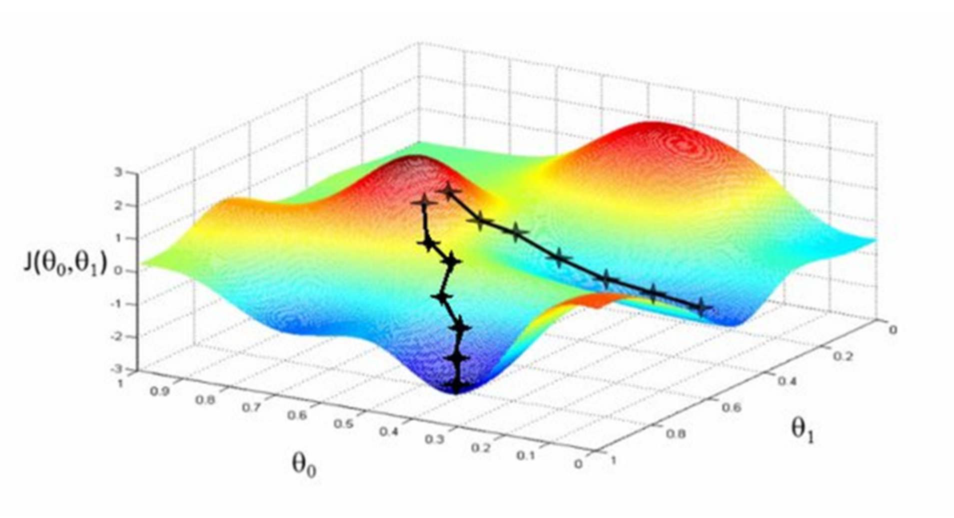

Figure 1

If we iteratively follow the direction of the -gradient (the greatest descent), we will finally get at the minimal value of the graph. From this graph, we aim to reach the point that has the lowest error (or minimum).

2.1.1 An illustration for comprehending gradient descent

A fictitious situation can be used to demonstrate the fundamental intuition behind gradient descent. A person who is trapped in the mountains is attempting to descend (i.e., trying to find the global minimum). There is so much fog that visibility is really limited. Since the route down the mountain is thereby hidden, they must employ local knowledge to determine the absolute minimum. They can employ the gradient descending technique, which entails assessing how steep the hill is at their current location before moving in that direction (i.e., downhill). They would move in the direction of the steepest ascent if they were attempting to locate the mountain's peak (or the maximum) (i.e., uphill). They might finally descend the mountain using this technique, or they would become trapped in a hole (local minimum or saddle point), such as a mountain lake. Assume, however, that the hill's steepness cannot be determined by mere observation but rather requires a complex measuring device, which the person just so happens to be holding at the time. They should use the device as little as possible if they wished to descend the mountain before dusk because it takes some time to measure how steep the hill is. The challenge then becomes determining how frequently they should measure the hill's steepness in order to stay on course. The algorithm in this comparison is represented by the person, and the algorithm's exploration of various parameter values is represented by the path travelled down the mountain. The slope of the function at that particular point is represented by the steepness of the hill. Differentiation is the tool used to measure steepness. They move in a direction that is consistent with the function's gradient at that particular location. The step size is determined by how far they move before taking another measurement.

2.1.2 Steps for determining the step size and descent direction

Finding a proper value for γ is a crucial practical issue since choosing a step size that is too small would hinder convergence and using a step size that is too big would cause divergence. While it may seem counter-intuitive to use a direction other than the steepest descending direction, the notion is that the smaller slope may be made up for by being sustained over a much greater distance.

Mathematical analysis of this can be done by considering a direction $\mathbf{p}_n$ and step size γ_n as well as the more generic update:

$$\mathbf{a}_{n+1} = \mathbf{a}_n - \gamma_n \, \mathbf{p}_n$$

It takes some consideration to get appropriate values for $\mathbf{p}_n$ and γ_n. We would first like the update direction to be downward. Mathematically, letting θ_n denote the angle between $-\nabla F(\mathbf{a_n})$ and $\mathbf{p}_n$, this requires that $\cos\theta_n > 0$. We require more details regarding the goal function that we are optimising before we can say more. We may demonstrate the following under the somewhat weak assumption that F is continuously differentiable:

$$F(\mathbf{a}_{n+1}) \leq F(\mathbf{a}_n) - \gamma_n \|\nabla F(\mathbf{a}_n)\|_2 \|\mathbf{p}_n\|_2 \left[\cos\theta_n - \max_{t\in[0,1]} \frac{\|\nabla F(\mathbf{a}_n - t\gamma_n \mathbf{p}_n) - \nabla F(\mathbf{a}_n)\|_2}{\|\nabla F(\mathbf{a}_n)\|_2}\right]$$

This inequality suggests that a trade-off between the two terms in square brackets determines the degree to which we can be certain the function F is lowered. The angle between the direction of descent and the negative gradient is measured by the first term enclosed in square brackets. The second term gauges the rate at which the gradient shifts along the path of descent.

In theory, inequality could be improved upon $\mathbf{p}_n$ and γ_n to choose an optimal step size and direction. The problem is that evaluating the second term in square brackets requires evaluating $\nabla F(\mathbf{a}_n - t\gamma_n\mathbf{p}_n)$, and extra gradient evaluations are generally expensive and undesirable. Some ways around this problem are:

1. Neglect the advantages of a clever descent path by setting $\mathbf{p}_n = \nabla F(\mathbf{a_n})$ and use line search to find a suitable step-size γ_n, such as one that satisfies the Wolfe conditions. Backtracking line search is a more practical method of selecting learning rates that has both positive experimental and theoretical results. Note that one does not need to choose $\mathbf{p}_n$ to be the gradient; any direction that has positive intersection product with the gradient will result in a reduction of the function value (for a sufficiently small value of γ_n).

2. Under the assumption that F is twice-differentiable, use its Hessian $\nabla^2 F$ to estimate $\|\nabla F(\mathbf{a}_n - t\gamma_n\mathbf{p}_n) - \nabla F(\mathbf{a}_n)\|_2 \approx \|t\gamma_n \nabla^2 F(\mathbf{a}_n)\mathbf{p}_n\|$. Then choose $\mathbf{p}_n$ and γ_n by optimising inequality.

3. Under the assumption that ∇F is Lipschitz, use its Lipschitz constant L to bound $\|\nabla F(\mathbf{a}_n - t\gamma_n\mathbf{p}_n) - \nabla F(\mathbf{a}_n)\|_2 \leq Lt\gamma_n\|\mathbf{p}_n\|$. Then choose $\mathbf{p}_n$ and γ_n by optimising inequality.

4. Generate a customized model of $\max_{t\in[0,1]} \frac{\|\nabla F(\mathbf{a}_n - t\gamma_n \mathbf{p}_n) - \nabla F(\mathbf{a}_n)\|_2}{\|\nabla F(\mathbf{a}_n)\|_2}$ for F. Then select $\mathbf{p}_n$ and γ_n by optimising inequality.

5. More sophisticated strategies might be feasible under stronger assumptions about the function F, such as convexity.

Most of the time, convergence to a local minimum can be ensured by following one of the recipes listed above. Gradient descent can converge to the global solution when the function F is convex because all local minima are also global minima.

2.2 Finance

The study of economics and finance benefits greatly from the study of multi-variable calculus. Multi-variable calculus helps us to examine the many input values that almost all models depend on. For instance, the Black-Scholes Equation for pricing derivatives was developed in mathematical finance as a result of research into the randomised (stochastic) processes involved in the price of stocks. Similar to this, Langragian multipliers can be used to solve straightforward studies of multi-variable economic utility functions bound by various types of income/budget constraints.

The Black-Scholes equation, or PDE, governs the price evolution of a European call or European put under the Black-Scholes model in mathematical finance. The word could broadly apply to a similar PDE that can be derived for a number of options, or it could refer to derivatives more generally.

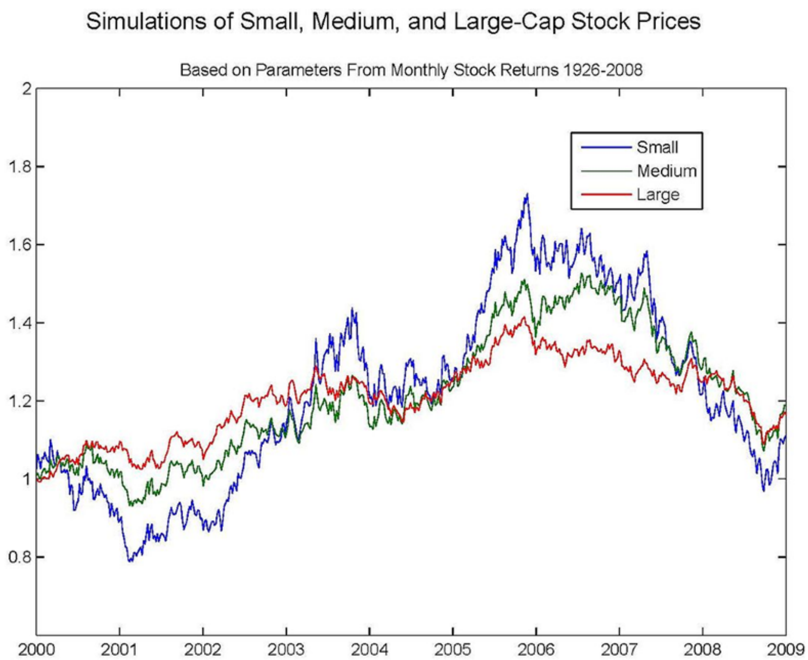

Figure 2: Simulated geometric Brownian motions with parameters from market data

2.2.1 Financial analysis of the PDE via Black-Scholes

The common derivation is based on the equation's concrete meaning, which is widely used by practitioners, and is provided in the following subsection. It is possible to write the formula as follows:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV - rS\frac{\partial V}{\partial S}$$

The terms that make up the left side of the equation are theta, the change of the derivative value with respect to time, and gamma, the convexity of the second spatial derivative with respect to the underlying value. A long position in the derivative and a short position consisting of the risk-free return $\partial V/\partial S$shares of the underlying are displayed on the right side.

According to Black and Scholes' discovery, the portfolio on the right represents a risk-free portfolio. Since theta and a term containing gamma are products, the risk less return over any infinitesimal time interval can be defined as such, according to the equation. Options frequently have negative theta values, which represent the value loss brought on by having less time to exercise an option (for a European call on an underlying without dividends, it is always negative). Gamma, which is typically positive, stands for the advantages of preserving the option. In order to generate a return at the risk less rate, the equation states that the gain from the gamma term and the loss from the theta term must balance each other out over any minuscule time interval. From the viewpoint of the option issuer, such as an investment bank, the gamma term refers to the cost of hedging the option.

2.2.2 Development of the Black–Scholes PDE

Hull's Options, Futures, and Other Derivatives has the following derivation. The classic justification in the original Black-Scholes article serves as the foundation for it as well.

According to the aforementioned model assumptions, the underlying asset's price moves in a geometric Brownian motion, which is

$\frac{dS}{S} = \mu\, dt + \sigma\, dW$

where W is a stochastic variable. Note that the sole variable in the stock's price history is W, and as a result, its infinitesimally small increase dW. W(t) is a process that, intuitively, "wiggles up and down" in such a random manner that its anticipated change over any time interval is 0. A simple random walk is an effective discrete analogue for W (and its variation over time T is equal to T as well). In light of the aforementioned calculation, the expected value of the stock's infinitesimal rate of return is μ dt and the variance is $\sigma^2 dt$.

An option's payout at maturity is known as V(S,T). We must understand how V changes as a function of S and t in order to determine its value earlier in time. For two variables, we have Itô's lemma such that

$$dV = \left(\mu S \tfrac{\partial V}{\partial S} + \tfrac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \tfrac{\partial^2 V}{\partial S^2}\right) dt + \sigma S \tfrac{\partial V}{\partial S}\, dW$$

Now consider a certain portfolio, called the delta-hedge portfolio, consisting of being short one option and long $\partial V / \partial S$ shares at time t. The value of these holdings is

$$\Pi = -V + \int \frac{\partial V}{\partial S} dS$$

Over the time period $[t, t + \Delta t]$ the total profit or loss from changes in the values of the holdings is:

$\Delta\Pi = -\Delta V + \tfrac{\partial V}{\partial S}\, \Delta S$

Now discretize the equations for dS/S and dV by replacing differentials with deltas:

$\Delta S = \mu S\, \Delta t + \sigma S\, \Delta W$

$$\Delta V = \left(\mu S \tfrac{\partial V}{\partial S} + \tfrac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \tfrac{\partial^2 V}{\partial S^2}\right) \Delta t + \sigma S \tfrac{\partial V}{\partial S}\, \Delta W$$

and appropriately substitute them into the expression for $\Delta\Pi$:

$$\Delta\Pi = \left(-\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)\Delta t$$

Notice that the ΔW term has vanished. As a result, uncertainty has been removed, making the portfolio practically risk-free. Otherwise, there would be chances for arbitrage. The rate of return on this portfolio must be identical to the rate of return on any other riskless product. Assuming that r is the risk-free rate of return, we must have over the period of time [t,t+Δt].

$$r\Pi\,\Delta t = \Delta\Pi$$

If we now equate our two formulas for $\Delta\Pi$ we obtain:

$$\left(-\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)\Delta t = r\left(-V + S\frac{\partial V}{\partial S}\right)\Delta t$$

We get to the well-known Black-Scholes partial differential equation by simplifying:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S}\mathrm{rV} = 0$$

This second order partial differential equation holds under the Black-Scholes model assumptions for every kind of option, provided that its price function V is twice differentiable with respect to S and once with respect to t. The selection of the reward function at expiration and the proper boundary conditions will result in various pricing formulas for various options.

Alternative derivation Here is a different derivation that might be used when it is first unclear what the hedging portfolio should consist of.

The underlying stock price S(t) is expected to change according to a geometric Brownian motion in the Black-Scholes model, presuming we have chosen the risk-neutral probability measure:

$\frac{dS(t)}{S(t)} = \text{r dt} + \sigma dW(t)$

Any derivative on this underlying is a function of time t and the stock price at the moment since this stochastic differential equation (SDE) demonstrates that the stock price evolution is Markovian. The discounted derivative process $e^{-rt}V(t, S(t))$,then has an SDE, which should be a martingale, when Ito's lemma is applied. The drift term must be zero for that to hold, which implies the Black—Scholes PDE.

2.3 Electrical Engineering

Numerous electromagnetic applications of multi-variable calculus exist. The Maxwell's Equations are the fundamental rules of electromagnetic upon which everything is based. These equations relate to magnetic and electric fields (which are vector fields which need multi-variable calculus). Since vector fields are virtually always involved in other electromagnetic laws, multi-variable calculus is a useful tool. The movement of charge (electrodynamics), which frequently results in a changing electric field and frequently produces a changing magnetic field, is what makes circuits function. All of these phenomena are governed by multi-variable calculus equations, namely Maxwell's Equations. In order to determine antenna performance, electromagnetic compatibility, radar cross section, and electromagnetic wave propagation outside of free space, it often entails computing approximations to solutions to Maxwell's equations using computer programs. Computer programs for antenna modelling, which calculate the radiation pattern and electrical characteristics of radio antennas and are frequently used to design antennas for particular applications, are a significant sub-field.

Due to the numerous irregular geometries present in actual devices, many real-world electromagnetic problems, such as electromagnetic scattering, electromagnetic radiation, modelling of wave-guides, etc., cannot be analytically calculated. The inability to construct closed form solutions of the Maxwell's equations under diverse boundary conditions and media constitutive relations can be addressed by computational numerical approaches. The design and modelling of antenna, radar, satellite, and other communication systems, as well as nanophotonic devices and high-speed silicon electronics, medical imaging, and cell-phone antenna design, are all major areas where computational electromagnetics is crucial.

2.3.1 Choice of methods

The right approach must be chosen when solving a problem since the incorrect approach can produce either inaccurate results or results that take too long to compute. Even for commercial tools, which frequently offer many solvers, the name of a technique does not always indicate how it is used.

2.3.2 Maxwell's equations in hyperbolic PDE form

A hyperbolic system of partial differential equations can be used to represent Maxwell's equations. Access to potent methods for numerical solutions is made possible by this.

The magnetic field must be parallel to the z-axis in order for the electric field to be parallel to the (x,y) plane for the waves to propagate in the (x,y)-plane. Transverse magnetic (TM) wave is the name of the wave.

The interplay of magnetic fields $\vec{B}$ and electric fields $\vec{E}$ over time is determined by Maxwell's equations. We consider the gradient, divergence, and curl to take partial derivatives in space, just like the Navier-Stokes equations (and not time t). Maxwell's system can therefore be expressed as follows (in "strong" form):

Gauss's law for electric fields: $\Delta.\vec{E} = \frac{\rho}{\epsilon_0}$

Gauss's law for magnetism: $\Delta.\vec{B} = 0$

Faraday's law: $\Delta \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$

Amp‘ere's law: $\Delta \times \vec{B} = \mu_0(\vec{J} + \epsilon_0 \frac{\partial \vec{E}}{\partial t})$

where ϵ_0 and μ_0 are physical constants and $\vec{J}$ encodes the density of electrical current. Just like the NavierStokes equations, Maxwell's equations related derivatives of physical quantities in time t to their derivatives in space.

By employing the differential version and the proper application of the Gauss and Stokes formula, Maxwell's equations can be written down with potentially time-dependent surfaces and volumes.

1. $\oint_{\partial\Omega}$ is a surface integral over the boundary surface $\partial\Omega$ with the loop indicating the surface is closed

2. $\iiint_{\Omega}$ is a volume integral over the volume ,

3. $\oint_{\partial\Sigma}$ is a line integral around the boundary curve $\partial\Sigma$, with the loop indicating the curve is closed.

4. $\iint_{\Sigma}$ is a surface integral over the surface Σ The total electric charge Q enclosed in ω is the volume integral over ω of the charge density ρ

$Q = \iiint_{\Omega} \rho \, \mathrm{d}V$, where dV is the volume element.

5. The net electric current I is the surface integral of the electric current density J passing through a fixed surface, Σ :

$I = \iint_{\Sigma} \mathbf{J} \cdot \mathrm{d}\mathbf{S}$,

where dS denotes the differential vector element of surface area S, normal to surface Σ.

In 2D and no polarization terms present, Maxwell's equations can then be formulated as:

$\frac{\partial}{\partial t}\bar{u} + A\frac{\partial}{\partial x}\bar{u} + B\frac{\partial}{\partial y}\bar{u} + C\bar{u} = \bar{g}$

where u, A, B, and C are defined as

$$\bar{u} = \begin{pmatrix} E_x \\ E_y \\ H_z \end{pmatrix},$$

$$A=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\epsilon} \\ 0 & \frac{1}{\mu} & 0 \end{pmatrix},$$

$$B=\begin{pmatrix} 0 & 0 & \frac{-1}{\epsilon} \\ 0 & 0 & 0 \\ \frac{-1}{\mu} & 0 & 0 \end{pmatrix},$$

$$C=\begin{pmatrix} \frac{\sigma}{\epsilon} & 0 & 0 \\ 0 & \frac{\sigma}{\epsilon} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The forcing function in this representation, $\bar{g}$ is located in the same region as $\bar{u}$. It can be used to define an optimization constraint or to specify an externally imposed field. As formulated above:

$$\bar{g} = \begin{pmatrix} E_{x,\text{constraint}} \\ E_{y,\text{constraint}} \\ H_{z,\text{constraint}} \end{pmatrix}.$$

$\bar{g}$ may also be explicitly declared equal to zero in order to identify a characteristic solution, which is frequently the first step in a process to find the specific inhomogeneous solution, or to simplify certain issues.

3 Conclusion

In this report, we have seen 3 applications of multi-variable calculus in Computer Science, Finance and Electrical Engineering domains in context of gradient descent, Black Scholes model and Maxwell equations. Apart from these, there are several other applications of multi-variable calculus in these domains as well as in various other domains. The main objective of this report was to convey just a few significant applications of multi-variable calculus in some domains, which is hopefully successfully done.

4 References

- https://en.wikipedia.org/wiki/Multivariable$_calculus : \, : text = Multivariate$
- https://collegedunia.com/exams/multivariable-calculus-steps-applications-and-examples-mathematics-articleid-5479
- https://en.wikipedia.org/wiki/Gradient$_descent$
- https://graphics.stanford.edu/courses/cs205a-13-fall/assets/notes/chapter14.pdf
- https://en.wikipedia.org/wiki/Black
- https://en.wikipedia.org/wiki/Maxwell