GEORGIAN AMERICAN UNIVERSITY

BUSINESS RESEARCH CENTER

BUSINESS ADMINISTRATION

Research Papers

Georgian American University (GAU)

Business Research Center

Conference

Applications of Stochastic Processes and Mathematical Statistics to Financial Economics and Social Sciences VII

November 24 - 25, 2022,
GAU, 10 Merab Aleksidze Str., 0160,
Tbilisi, Georgia

Conference Materials

Content

The Future of Neobanking

Dr. Besarion Abuladze, PhD, MBA
Professor of Georgian American University

Today we are witnessing the digital revolution, which in the coming decade will totally disrupt the Global Banking Sector. Neobanks are part of this revolution.

To be more specific, "change" is the keyword today - the change in the lifestyle of Consumers. The lifestyle changes were always occurring due to demographics, technological developments, fluctuating financial conditions of consumers and many other factors. However, some additional factors have recently accelerated this process, such as the fear for health due to COVID-19 and the consequent necessity for isolation.

In any case, the change in lifestyle is affecting the purchasing habits of Consumers. The change in purchasing habits of Consumers is manifested at each logical step of purchasing behavior, such as: the ways of searching for and getting the information, evaluating alternatives, making purchase decision and engaging into the purchasing process itself.

To be successful, the changes in Consumer purchasing behavior should be matched with the adequate Customer Value Proposition by the banking service providers, including Neobanks.

But, does the value proposition of the financial services providers (including Neobanks) match the changing purchasing habits of Consumers?

To answer this question, we have to start with the definition: "A neobank (also known as an online bank, internet-only bank, virtual bank or digital bank) is a type of direct bank that operates exclusively online *without traditional physical branch networks*" [1]. However this is the definition made from the perspective of a service provider. Now let's think from the perspective of a customer. The customer of a Neobank receives banking (or neobanking) services wholly online without visiting a branch.

Now, if a reader asks herself or himself - does it matter for me, as a customer, who provides to me the online banking services that I need: a truly online bank, or a bank that has some branches, or someone that even does not have a full banking license (e.g. payment service providers)? The answer will be: no, it does not matter for me, as long as I receive the service I need.

That's why in this article we are referring to Neobanking (please refer to the title of the present article) as to a wholly online banking service provided by different institutions, such as Fintechs and even traditional banks.

The forecast is that the market size of the Neobanks (in its traditional definition) will "grow at an annual average rate (CAGR) of 53.4 percent until 2030, reaching a value of 2.05 trillion U.S. dollars that year" [2]. However, if we use a more broad definition of Neobanking also including the digital banking services, then the market size will grow by around 3.6 percent CAGR reaching more than 13 trillion U.S. dollars in 2032 [3].

There are approximately 30,000 Fintech companies in the World (October 2022) [4] with the total value of investments into the Fintech Sector exceeding 1 trillion U.S. dollars [5]. The extreme competition within the industry is intensified because the traditional banks also transfer their products and services into digital channels. Take the example of two leading banks in Georgia: the retail offloading ratio of TBC Bank was 97% in 2021 [6], while the share of retail transactions through digital channels of Bank of Georgia was 96,1% in the same year [7].

The ways how to overcome the above mentioned extreme competition that is taking place in the Financial (Fintech) Sector, can be found in the Marketing Principles. Namely, back in 1990's article – "Four P's Passe; C-words Take Over" [8], Bob Lauterborn introduces the 4 C's (Customer, Cost, Convenience, Communication), which replace the 4 P's of the original marketing mix concept, as shown in Fig. 1 below:

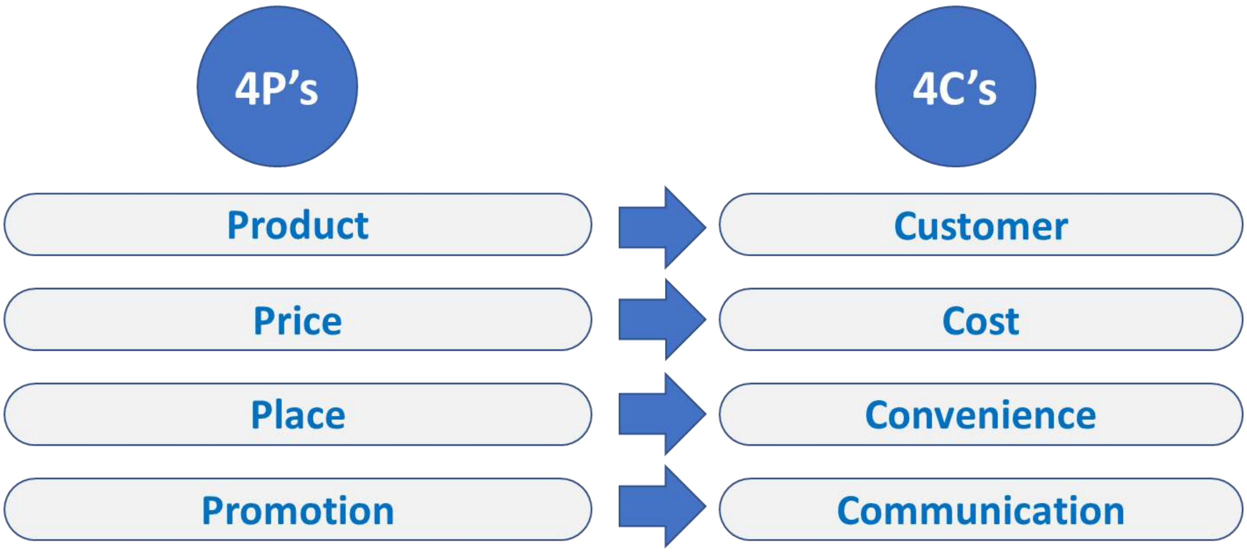


Fig. 1 Four P's of marketing mix is replaced by four C's

This shift from 4 P's to 4C's shows how the digital revolution is affecting the established principles: "product" is replaced by "customer", because mass customization in the digital world allows companies to deliver values to customers on an individualized way; "price" is replaced by "cost", because the marginal cost of distributing a digital product is zero; "promotion" is replaced by "communication", because the digital channels allow for direct communication to end users in a cost-effective manner.

The special comment should be devoted to "place", which is now replaced by "convenience". As it was suggested by Bob Lauterborn in his original article: "Forget place. Think convenience to buy. People don't have to go anyplace any more, in this era of catalogs, credit cards and phones in every room" [8]. This was said in the era of traditional landline phones, before emergence of mobile phones as a means of communication and a way before the smartphones emerged as platforms for receiving all kinds of digital products and services.

Today it is easy to shop, to pay, to order food or products, to book, to travel, to receive various kinds of digital products and services by using a mobile phone, all done in one place. Is it convenient? Definitely it is. That's how "convenience" has replaced "place". To summarize, the shift in marketing mix shown in Fig. 1 is caused by the changes in consumer behavior, which itself is affected by technological changes.

In order to demonstrate how technological and lifestyle changes may affect the consumer behavior, we can resort to historical parallels. The period after the WWII, especially the mid-1950s, experienced the increased car ownerships in the USA, which increased the mobility of shoppers. With the increased mobility, it became a lifestyle to combine shopping, entertainment and leisure in one trip as it was convenient for consumers. The evolved lifestyle has manifested itself in the changed consumer behavior. The retailing sector responded to this change with the emergence of shopping malls by means of providing all services in one place in a manner convenient for consumers.

Basically, what had happened in the mid of the last century in the USA, is nowadays explained by a so called “Blue Ocean Strategy” [9]. It defines the process of “simultaneous pursuit of differentiation and low cost to open up a new market space and create new demand. It is about creating and capturing uncontested market space, thereby making the competition irrelevant. It is based on the view that market boundaries and industry structure are not a given and can be reconstructed by the actions and beliefs of industry players” [10].

The Blue Ocean strategy is essentially the adaptation of a business to the shift occurring in consumer purchasing behavior. Although this theory was invented after 50 years from the massive occurrence of the shopping malls in the USA, it explains the rationale behind the adoption of the shopping malls as shown in Fig. 2. At that time the shopping malls differentiated themselves in the eyes of consumers (offering all services in one place in a convenient manner) and restructured the cost base for retailers (offering lower cost rentals to retailers compared to downtown locations).

Fig. 2 Shift in consumer behavior resulted in the emergence of Shopping Malls in the USA

Nowadays, it is *convenient* for consumers to use their mobile phones for receiving many different services - *all in one place* (that’s how the “convenience” has replaced the “place” in the marketing mix). It only remains up to the businesses to adapt to the changed consumer behavior in a most convenient way for customers as it can be seen from Fig. 3.

Fig. 3 Modern-day shift in consumer behavior drives the emergence of Super Apps

Now we are in the position to answer the main question posed in this article: *does the value proposition of the financial services providers (including Neobanks) match the changing purchasing habits of Consumers?* If we agree that the value for a modern-day consumer is "convenience" by means of receiving virtually all services in one place (i.e. in a mobile phone), then the financial service providers should respond by creating the respective ecosystems. Is this the case?

The answer is No. Today, the financial services industry is fragmented into digital payments, loyalty platforms and neobanking/digital banking services to mention just a few. In addition, there are numerous companies offering E-commerce services/marketplaces, which increase the degree of defragmentation from the consumers' perspective. All these services are facing some problems when run in isolation as shown in Fig. 4 below:

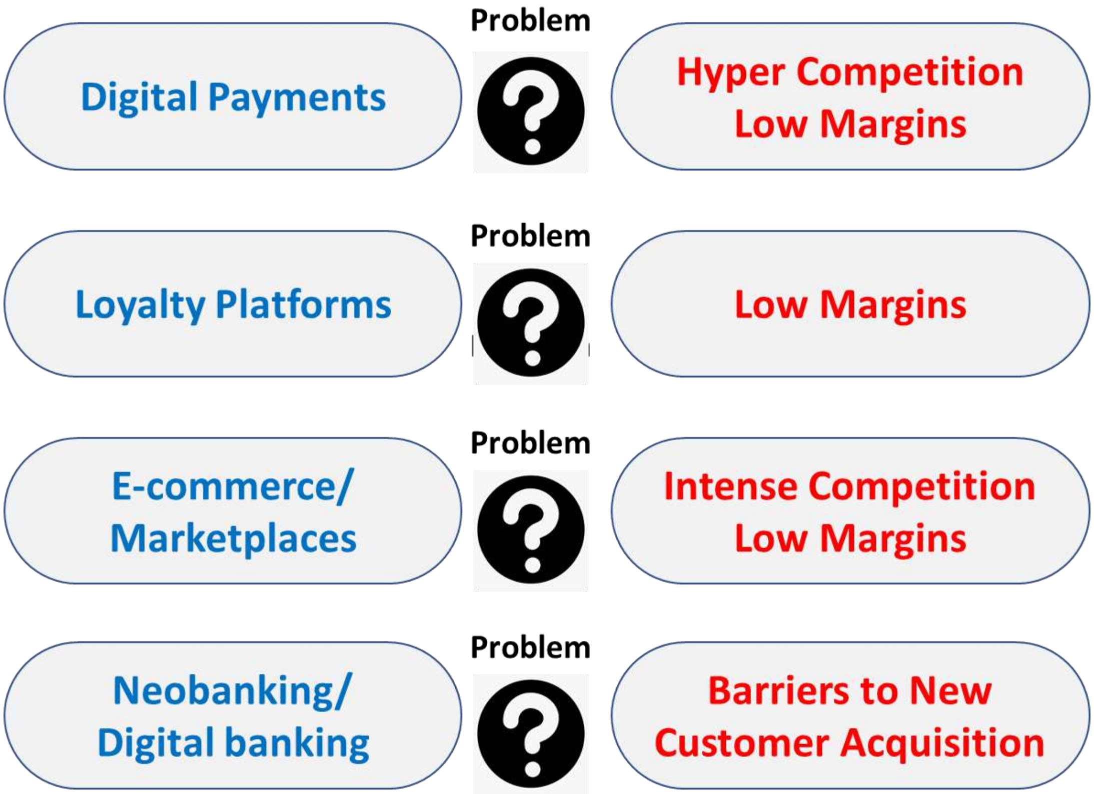


Fig. 4 Problems facing the individual industry players if run in isolation

It then follows that the individual industry players should think about creating more complex ecosystems. With such approach they will enhance the value proposition for consumers by offering many services in one place (convenience) and at the same time will overcome some of the above mentioned problems by means of achieving synergy effects and the economies of scale.

There are many options of incorporating different services into a single ecosystem of services. For example, consider loyalty platforms and digital banking. From the above statement of problems (refer to Fig. 4 above) let us choose two main dimensions: profit potential per customer and potential of new customer acquisition. As a next step, let us compare loyalty platforms and digital banking services with each other according to these two dimensions. The comparison is shown in Fig. 5:

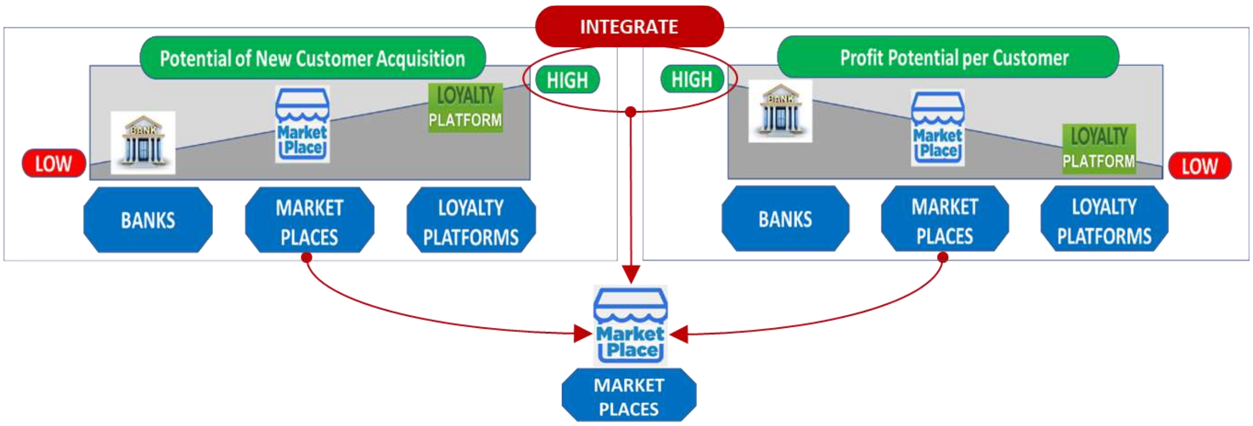


Fig. 5 Comparison of loyalty platforms and digital banking services

It can be seen from Fig. 5 that loyalty platforms have the highest potential of new customer acquisition because of the attractiveness of such programs to many consumers. At the same time the loyalty platforms have the lowest margins due to its nature. On the other hand (refer to Fig. 5), we can see that neobanks/digital banks have the lowest potential of new customer acquisition due to high competition and due to the barriers established by the traditional banks. At the same time the banks have the highest profit potential due to a high Customer Lifetime Value (CLV).

Therefore, the solution will be to unify both into a single ecosystem. But in order to originate the necessary transactions for banks, one will need to incorporate an additional component into the proposed ecosystem. Such component may be a marketplace as shown in Fig.5.

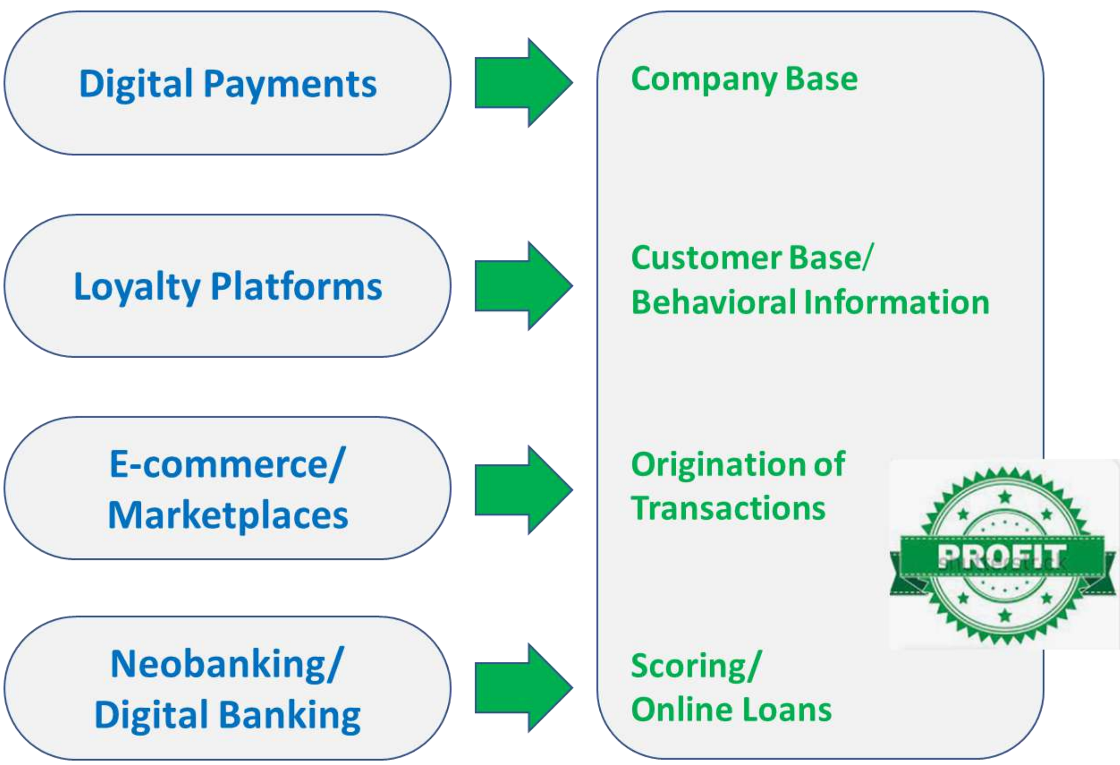


Fig. 6 The resulting ecosystem

The resulting ecosystem is shown in Fig.6. The business logic behind such ecosystem is the following: digital payments functionality will generate a large base of participating companies, as these companies look for reducing the transaction fees. Loyalty platforms will generate a large customer base and provide customer behavioral information to the participating banks (with the consent of customers). Banks will score the customer behavior information and will embed their offerings into the marketplace. As a result the banks will be able to issue online loans for consumers shopping at marketplace. The customers of the ecosystem will receive all the above mentioned services in one place (convenience). The resulting ecosystem will be profitable due to high CLV of banking products and services.

Summary

In order to summarize the future development of neobanking in just a few words, it can be stated: neobanking is the future. A lot of new financial ecosystems will emerge, some of them in partnership with the existing banking institutions. The examples of such initiatives can be seen everywhere: Visa and Mastercard are becoming digital, Apple and Amazon are incorporating the financial services, Paypal is extending its services into the loyalty industry, even Twitter and Google are seeing themselves as payment service providers. It only remains to see how these neobanking ecosystems will be reshaped in the coming decade.

Literature

1. https://en.wikipedia.org/wiki/Neobank
2. https://www.statista.com/statistics/1228241/neobanks-global-market-size/
3. https://www.gminsights.com/industry-analysis/digital-banking-market
4. https://explodingtopics.com/blog/fintech-stats
5. https://www.statista.com/statistics/719385/investments-into-fintech-companies-globally/
6. https://www.tbcbank.ge/web/documents/10184/616784/Pillar+3+Report+_2021_+26+April.pdf/828379bb-2639-4431-a8ed-8648837eb2df
7. https://bankofgeorgiagroup.com/storage/reports/ARA2021_for%20web.pdf
8. Lauterborn, B. “New marketing litany; four P’s passe; C-words take over”, 1990, Advertising Age; Vol 41; pp. 26, http://rlauterborn.com/pubs/pdfs/4_Cs.pdf
9. Kim, W. C. and R. Mauborgne, “Blue ocean strategy”, 2004, Harvard Business Review (October), pp. 76-84, https://hbr.org/2004/10/blue-ocean-strategy
10. https://www.blueoceanstrategy.com/what-is-blue-ocean-strategy/

The importance of venture capital in innovative investment projects

Dali Magrakvelidze
Georgian Technical University
E-mail: d.magrakvelidze@gtu.ge

Abstract

Innovation makes it possible to produce more products with less materials and resources, which in turn leads to economic growth. The main problem of financing innovative projects is the high risk of returns and long payback. Most of these projects do not have enough guarantee funds, their resources are limited, and only their own ideas and technologies are the backbone. Due to the high risk of innovative projects, it is necessary to use venture capital to finance them.

Key words: Venture capital, venture business, innovation, "valley of death".

* * * * *

In the modern world, as the population grows, the role of technological innovations in meeting human needs increases, as they change the world economy and contribute to the economic growth of countries.

Innovation and entrepreneurship are the kernels of a capitalist economy. New businesses, however, are often highly-risky and cost-intensive ventures. As a result, external capital is often sought to spread the risk of failure. In return for taking on this risk through investment, investors in new companies are able to obtain equity and voting rights for cents on the potential dollar. Venture capital, therefore, allows startups to get off the ground and founders to fulfill their vision.

Venture business involves financing new ideas, progressive scientific and technical developments and bringing them down to a suitable level for sale, i.e. commercialization. Venture business requires a lot of knowledge, a lot of money, and a lot of guts, but if successful, it can be hugely profitable. This type of business does not actually exist in our country, because we do not have the experience of working with new technologies and risky investments, as well as the financial infrastructure.

Venture capital (VC) is a form of private equity and a type of financing that investors provide to startup companies and small businesses that are believed to have long-term growth potential. Venture capital generally comes from well-off investors, investment banks, and any other financial institutions.

However, it does not always take a monetary form; it can also be provided in the form of technical or managerial expertise. Venture capital is typically allocated to small companies with exceptional growth potential, or to companies that have grown quickly and appear poised to continue to expand. Venture capital funds manage pooled investments in high-growth opportunities in startups and other early-stage firms and are typically only open to accredited investors.

One important difference between venture capital and other private equity deals, however, is that venture capital tends to focus on emerging companies seeking substantial funds for the first time, while private equity tends to fund larger, more established companies that are seeking an equity infusion or a chance for company founders to transfer some of their ownership stakes.

Venture capital is a subset of private equity (PE). While the roots of PE can be traced back to the 19th century, venture capital only developed as an industry after the Second World War.

Harvard Business School professor Georges Doriot is generally considered the "Father of Venture Capital." He started the American Research and Development Corporation (ARD) in 1946 and raised a $3.5 million fund to invest in companies that commercialized technologies developed during WWII. ARDC's first investment was in a company that had ambitions to use x-ray technology for cancer treatment. The $200,000 that Doriot invested turned into $1.8 million when the company went public in 1955. [3]

A series of regulatory innovations further helped popularize venture capital as a funding avenue.

The first one was a change in the Small Business Investment Act (SBIC) in 1958. It boosted the venture capital industry by providing tax breaks to investors. In 1978, the Revenue Act was amended to reduce the capital gains tax from 49% to 28%.

Then, in 1979, a change in the Employee Retirement Income Security Act (ERISA) allowed pension funds to invest up to 10% of their assets in small or new businesses. This move led to a flood of investments from rich pension funds.

The capital gains tax was further reduced to 20% in 1981.

These three developments catalyzed growth in venture capital and the 1980s turned into a boom period for venture capital, with funding levels reaching $4.9 billion in 1987. The dot-com boom also brought the industry into sharp focus as venture capitalists chased quick returns from highly-valued Internet companies. According to some estimates, funding levels during that period went as high as $30 billion. But the promised returns did not materialize as several publicly-listed Internet companies with high valuations crashed and burned their way to bankruptcy. [5]

The 2008 financial crisis was a hit to the venture capital industry because institutional investors, who had become an important source of funds, tightened their purse strings. The emergence of startups that are valued at more than a billion dollars, has attracted a diverse set of players to the industry. Sovereign funds and notable private equity firms have joined the hordes of investors seeking return multiples in a low-interest-rate environment and participated in large ticket deals. Their entry has resulted in changes to the venture capital ecosystem. Arthur Rock, an investment banker at Hayden, Stone & Co. in New York City helped facilitate that deal and subsequently started one of the first VC firms in Silicon Valley.

Venture Capital has some advantages and disadvantages. Advantages are:

Business expertise. Aside from the financial backing obtaining venture capital financing can a start-up or young business with a valuable source of guidance and consultation. This can help with a variety of business decisions, including financial management and human resource management.

Additional resources. In a number of critical areas, including legal, tax and personnel matters, a VC firm can provide active support, all the more important at a key benefits.

Connections. Venture capitalists are typically well connected in the business community. Using these connections can have huge benefits.

The main disadvantages are: Loss of control and minority ownership status.

Venture capital provides funding to new businesses that do not have access to stock markets and do not have enough cash flow to take debts. This arrangement can be mutually beneficial: businesses get the capital they need to bootstrap their operations, and investors gain equity in promising companies.

There are also other benefits to a VC investment. In addition to investment capital, VCs often provide mentoring services to help new companies establish themselves, and provide networking services to help them find talent and advisors. A strong VC backing can be leveraged into further investments.

On the other hand, a business that accepts VC support can lose creative control l over its future direction. VC investors are likely to demand a large share of company equity, and they may start making demands of the company's management as well. Many VCs are only seeking to make a fast, high-return payoff and may pressure the company for a quick exit.

Venture capital can be broadly divided according to the growth stage of the company receiving the investment. Generally speaking, the younger a company is, the greater the risk for investors.

The stages of VC investment are: Pre-Seed: This is the earliest stage of business development when the founders try to turn an idea into a concrete business plan. They may enroll in a business accelerator to secure early funding and mentorship; Seed Funding: This is the point where a new business seeks to launch its first product. Since there are no revenue streams yet, the company will need VCs to fund all of its operations; Early-Stage funding: Once a business has developed a product, it will need additional capital to ramp up production and sales before it can become self-funding. The business will then need one or more funding rounds, typically denoted incrementally as Series A, Series B, etc.

The main problem of financing innovative projects is the high risk of returns and long payback. Most of these projects do not have enough guarantee funds, their resources are limited, and only their own ideas and technologies are the backbone. The problem of financing such enterprises has received the concept of "valley of death" in economics. This is manifested in the fact that between the product project part and its launch on the market for a long time there are problems in terms of financing, which lead to large cash gaps and, as a result, insolvency, which threatens the existence of the project. The cause of the "valley of death" is the different goals of investors and businessmen (developers), the former strive for quick profit, and the latter are focused on obtaining scientific results.

For small businesses, or for up-and-coming businesses in emerging industries, venture capital is generally provided by high net worth individuals (HNWIs)—also often known as “angel investors"—and venture capital firms. The National Venture Capital Association (NVCA) is an organization composed of hundreds of venture capital firms that offer to fund innovative enterprises.

Common occurrence among angel investors is co-investing, in which one angel investor funds a venture alongside a trusted friend or associate, often another angel investor.

While both provide money to startup companies, venture capitalists are typically professional investors who invest in a broad portfolio of new companies and provide hands-on guidance and leverage their professional networks to help the new firm. Angel investors, on the other hand, tend to be wealthy individuals who like to invest in new companies more as a hobby or side-project and may not provide the same expert guidance. Angel investors also tend to invest first and are later followed by VCs.

Due to the industry's proximity to Silicon Valley, the overwhelming majority of deals financed by venture capitalists are in the technology industry—the internet, healthcare, computer hardware and services, and mobile and telecommunications. But other industries have also benefited from VC funding.

Venture capital is also no longer the preserve of elite firms. Institutional investors and established companies have also entered the fray. For example, tech behemoths Google and Intel have separate venture funds to invest in emerging technology. In 2019, Starbucks also announced a $100 million venture fund to invest in food startups.

Data from the NVCA and PitchBook indicate that venture-backed companies have attracted a record $330 billion in 2021, compared to the total of $166 billion seen in 2020—which was already a record. Large and late-stage investments remain the main drivers behind the strong performance: Mega-deals of $100 million or more have already hit a new high-water mark. [4]

Another noteworthy trend is the increasing number of deals with non-traditional VC investors, such as mutual funds, hedge funds, corporate investors, and crossover investors. Meanwhile, the share of angel investors has gotten more robust, hitting record highs, as well.

Late-stage financing has become more popular because institutional investors prefer to invest in less-risky ventures (as opposed to early-stage companies where the risk of failure is high).

Innovation and entrepreneurship are the kernels of a capitalist economy. New businesses, however, are often highly-risky and cost-intensive ventures. As a result, external capital is often sought to spread the risk of failure. In return for taking on this risk through investment, investors in new companies are able to obtain equity and voting rights for cents on the potential dollar. Venture capital, therefore, allows startups to get off the ground and founders to fulfill their vision.

New companies often don't make it, and that means early investors can lose all of the money that they put into it. A common rule of thumb is that for every 10 startups, three or four will fail completely. Another three or four either lose some money or just return the original investment, and one or two produce substantial returns.

Depending on the stage of the company, its prospects, how much is being invested, and the relationship between the investors and the founders, VCs will typically take between 25 and 50% of a new company's ownership.

References:

1. J. Lerner, R. Nanda, Venture Capital`s Role in Financing Innovation: What We Know and How Much We Still Need to Learn, Working Paper 20-131, Harvard Business School, 2020.

2. N. Khan, H. Qu, J. Qu, Ch. Wei, sh. Wang, Does Venture Capital Investment Spur Innovation? A Cross-Countries Analysis, SAGE Open January-March 2021: 1–13.

3. The Business History Conference. "The Rise and Fall of Venture Capital," Pages 6-8.

4. National Venture Capital Association. "Venture Monitor Q4 2021."

5. https://www.investopedia.com/terms/v/venturecapital.asp#citation-15

Modern theories of leadership and types of leaders

Ketevan Kiguradze
PhD Student, Business School, Georgian American University, Tbilisi, Georgia

ABSTRACT

Globalization has opened the borders of countries, due to which the competition between the companies intensified it became necessary to guide and manage not the usual a manager, but a globally minded manager or leader. Even today, the question of the phenomenon of leadership is controversial among scientists about, in particular, how should he be a leader, with qualities or skills, whether it is possible to learn leadership, etc.
Cross-Country Perceptions of Leadership Skills and Traits is different. There are different views on it in terms of vertical management hierarchy and network structuring. of business network structures clearly influence organizational culture he showed us about the leader and leadership.
In such conditions, when Georgia is also one of the members of the global world, in relation to the compatibility of Georgian culture, the issue of how it should be in Georgia should be studied the leader of a functioning company, what qualities and skills should have, what features of Georgian culture should he grasp and, finally, what should he be able to give for the success of the company "As a victim".
The world is changing at lightning speed and also at lightning speed public interest in leadership and leaders is growing. We all have a certain view of what leadership is, however the exact definition of the term is still difficult. Some scientists believe that leadership is the result of successful group dynamics, Others believe that leadership is determined by the grandiose efforts of a person. There is also an opinion that the leader is influenced by circumstances and necessity creates, however, according to another opinion, leadership a person with qualities is always a leader.

It is important to understand that the leadership style or type used by managers or leadership positions in a company always has consequences for workers, even if we do not realize it or confuse these consequences with the inner personality of each person. It is very important to clarify this, since leaders are agents who believe that they are in a privileged position to influence others for better or worse.

American businessman, founder of Microsoft company and CEO (Gates, Bill, n.d.)rightly notes: “We expect that in the next century, the leader will be the one who makes others authorized".

Transformational leadership is one of the most modern and popular leadership theories. It was founded in the 80s of the 20th century and is "part of the new leadership paradigm" (Peter Guy Northouse, 2010). The theory was based on the works of (Bass, B. M., & Riggio, R. E. ., 2006) (Burton Nanus, Warren G. Bennis, 2006) made important contributions to the development of the theory. According to (Bass, B. M., & Riggio, R. E. ., 2006), the popularity of transformational theory is likely due to its emphasis on intrinsic motivation and follower development. According to this theory, people at the level of change and uncertainty need inspiration and faith in themselves. In their (2001) analysis of articles published in the Quarterly Journal of Leadership, Lue and Gardner concluded that 1/3 of the studies were about transformational or charismatic leadership. Transformational leadership is one of the most comprehensive theories. This is a process that leads to changes and transformations in people. It deals with emotions, values, ethics, standards and long-term goals. Also, it includes evaluation of followers' motivation. meeting their needs and treating them with respect. According to this theory, leaders inspire and motivate followers to do great things and hold followers to high standards. According to this approach, the leader must understand and adapt his actions to the needs and motives of the followers. In transformational leadership, pseudo-transformational leadership is distinguished, which transforms in a negative way. Leaders who experience transformation in a negative way, are self-absorbed, focused on power and are carriers of distorted moral values are considered pseudo-transformational.

Transformational leadership also has some weaknesses. For example, it lacks conceptual clarity, another weakness relates to the measurement of transformational leadership. Some transformational factors are correlated with transactional and noninterventional leadership factors. It should also be noted that transformational leadership does not present clear assumptions about how leaders should act in specific situations. It focuses on ideals, inspiration, motivation, innovation and individual care.

During the same period, the researcher (Bass, 1985) proposed an even more sophisticated version of transformational leadership, which was based on the work of (House), but did not completely follow it. He believed that transformational leadership can be used in situations where the results are negative. He considered transactional and transformational leadership on the same line. (Avolio, 1999), in 1999, referred to transformational leadership as improving the performance of followers and realizing their full potential.

Transactional leadership includes all types of leadership that focus on agreement between leaders and followers. It encourages high performance through the use of rewards and punishments. When managers reward subordinates for good performance and, conversely, punish them for poor performance, they increase the motivation of subordinates to ensure the desired action.

Charismatic leadership is often compared to transformational leadership. As mentioned in the previous subsection, charisma was first defined by (Weber, 1974)), who describes it as follows: "It is a personal characteristic that gives a person superhuman, outstanding power, it is not available to everyone, it is of divine origin, and as a result we get a person who is perceived as a leader. and treated as a leader. Later this theory was developed by (House, 1971). the personality characteristics of a charismatic leader developed by (House), which include: dominance, a strong desire to influence others, self-confidence, and belief in one's own moral values.

According to House's charismatic leadership theory, its face-to-face outcome is the follower's trust in the leader's ideology. Recognition of the leader without any doubts or questions.

Authentic leadership is one of the newest areas in leadership research. The theory focuses on how "real" and how authentic leadership is. There are several definitions of authentic leadership that explain it from different perspectives, they are: intrapersonal - processes taking place inside the leader's personality, self-knowledge, self-regulation, and self-evaluation; Developing - leadership behavior that is formed from the positive psychological characteristics and high quality of the leader. This is what is formed in people throughout life. Interpersonal - is built on relationships and involves achieving interactions between leaders and followers. It is a two-way process, as leaders influence followers and vice versa.

Today, one of the most recognized approaches in the field of leadership research is (House). The theory of conformity of means and ends. The essence of this theory lies in what the leader does to motivate subordinates to achieve the group and organization's goal. 1. Effective leaders clearly define the goals that subordinates are trying to achieve by working; 2. They reward subordinates according to the work done and the goal achieved and 3. They make clear the path that leads to the work goal. According to this theory, the steps a leader should take to motivate subordinates depend on both the subordinates and the type of work performed. In the theory of compatibility of the goal and the means, four behaviors of the leader are distinguished: 1. directive behavior; 2. Supportive behaviors; 3. complicity behavior; 4. Achievement-oriented behavior. Therefore, leaders must decide for themselves which behavior to use during the task to be performed by the subordinate in order to motivate them to perform the task.

Leadership concepts address the factors that leaders consider when applying leadership styles and overseeing an individual team. These principles focus on the ideas and perceptions about the qualities that leaders should have and how they should perform in the role of leader. In addition, leadership concepts help professionals understand what kind of skills and character traits they need to develop to advance in leadership roles.

The concepts of leadership differ from leadership theories in several ways. For example, leadership concepts generally serve as a guide for professionals to use in shaping leadership styles, communicating with teams, and leading processes. Leadership theories typically focus on the idea of using different methodologies, styles, and techniques when leading a team. Leadership concepts include different styles, qualities, and principles of employee team management approaches. Essentially, leadership concepts are based on various theories of management, and these qualities serve as the standard for effective managers, leaders, and other positional leaders. In addition, leadership concepts form the basis of standard management style and behavior theories and often include traits such as personality and character, initiative, motivation, influence, decision-making ability.

According to the studies by (Kirkpatrick, S.A. and Locke, E.A, 1991)have identified six traits that distinguish leaders from others. These are: Attitude, motivation, honesty, self-confidence, cognitive abilities and knowledge of the case. They think people with similar traits can be born or acquired over a lifetime They are. These 6 traits are exactly the traits that leaders need. These qualities of a leader distinguish people from each other and therefore, these differences are an important part of the leadership process. Also, empirical research (Peter G Northouse, 2010) conducted in the 1990s has shown that with social intelligence comprehension of feelings, behavior, and thoughts related traits are important traits for an effective leader.

The discussion of leadership as a trait has aroused great interest among foreign scholars. (Bryman, A, 1992) and an analysis by (Ellinger, A. D, 1986)found that personal qualities are closely related to a person's views on leadership. Even according to (Kirkpatrick, S.A. and Locke, E.A, 1991) effective leaders with distinctive qualities in some respects there are people. It is interesting to note that in the 1990s leadership as a trait became the subject of special attention to those for researchers who are distinguished by visionary, or charismatic leadership. These are: (Bass, B. M. , 1985), (Burton Nanus, Warren G. Bennis, 1985), (David A. Nadler, 1989)

It will not be uninteresting to discuss all those theories and systematic research devoted to the topic of leadership and identifying the characteristics of prominent leaders, since, naturally, the achievements of great people lead to universal recognition.

Followers of **The Trait Theory** as the starting point of leadership consider the individual characteristics of a leader and try to study the leader through these characteristics. This approach known as The Great Man Theory. This theory was formulated by Scottish philosopher, writer and teacher Thomas Carlyle. The theory is based on two main assumptions: 1) Great leaders are born, they have certain qualities that allows a person to be guided, guided and to be promoted; 2) Great leaders emerge when their need arises exists. Significant research and agreement have been reached today that leadership comes from a combination of both theories - and more. As already mentioned, there is a wide selection of qualities and characteristics of a leader. The University of Santa Clara and the Tom Peters Group identified the following leadership characteristics: Honesty - Show sincerity, integrity and sincerity in all your actions. Deceptive behavior does not inspire confidence. Competent - base your actions on common sense and moral principles. Do not make decisions based on childish emotional desires or feelings. Prudent - Set goals and have a vision for the future. The vision should be owned by the entire organization. Effective leaders see what they want and how to get it. They usually choose priorities based on their core values. Inspiring - Show confidence in everything you do. By displaying mental, physical and spiritual endurance you inspire others to reach new heights. Take responsibility if necessary. Intelligent - Read, study and search for difficult tasks. Righteous Thinker - Treat all people fairly. Superstition is the enemy of justice. Show empathy for the feelings, values, interests, and well-being of others. Broad thinking - seek diversity. Brave - have the patience to achieve the goal, despite seemingly insurmountable obstacles. Exercise self-confidence in times of stress. Direct - Use common sense to make the right decisions at the right time. Imaginative - Make timely and appropriate changes in your thinking, plans and methods. Show creativity by thinking of new and better goals, ideas and problems. (John Whitehead, 2016)

Leadership theories study the qualities of effective leaders, including the qualities of effective and influential leaders, patterns of behavior, and actions. Leadership theories focus on explaining what makes good leaders by focusing on different behaviors and qualities that professionals can develop to become good leaders. While the concepts of leadership are qualities in themselves, leadership theories are the study and explanation of these qualities and their impact on professionals and their work environment.

Do not lose sight of the fact that there is no one type of leadership that is universally good or desirable. Organizations are characterized by their complexity, and therefore they can undergo changes at various and different rates, so you need to know how to adapt to momentary requirements; This also happens through his leadership, through the distribution of roles, etc. In all cases, the truth is that a leader is not a leader by tenure: leadership is exercised, it is not inherited, something you place in an organization chart.

Bibliography

Avolio, B. J. (1999). Full leadership development: Building the vital forces in organizations. *Thousand Oaks*.

Bass, B. M. . (1985). Leadership and performance beyond expectations. *New York: Free Press.*

Bass, B. M. (1985). Leadership and performance beyond expectations. . *New York: Free Press.*

Bass, B. M., & Riggio, R. E. . (2006). *Transformational leadership. .* Lawrence Erlbaum Associates Publishers.

Bryman, A. (1992). *Charisma and Leadership in Organization.* London: Sage Publications.

Burton Nanus, Warren G. Bennis. (1985). *Leaders: Strategies for Taking Charge.* Harper & Row.

Burton Nanus, Warren G. Bennis. (2006). Leaders: Strategies for Taking Charge. . *New York: HarperCollins Publisher*.

David A. Nadler. (1989). *The Management of Organizations: Strategies.* New York: Harper and Row.

Ellinger, A. D. (1986). A definitional examination of ‘coaching,’ ‘organization development,’ and ‘human resource development'. *Leadership & Organization Development Journal*.

Gates, Bill. (n.d.). Retrieved from forbes.com: https://www.forbes.com/sites/kevinkruse/2013/04/09/what-is-leadership/?sh=391dce1d5b90#2fbb706713e1;

House, R. (1971). *Path-Goal theory. .*

John Whitehead. (2016, 09 16). *https://johnkwhitehead.ca/three-theories-leadership/*. Retrieved from https://johnkwhitehead.ca/three-theories-leadership/.

Kirkpatrick, S.A. and Locke, E.A. (1991). *Leadership Do Traits Matter.* Lexington: Lexington Books.

Peter G Northouse. (2010). *leadership theory and practice.* Thousand Oaks, CA: Sage Publications .

Peter Guy Northouse. (2010). *Introduction to Leadership: Concepts and Practice.* SAGE 2007.

Weber. (1974). *Leadership style. .*

How to Compute the Gradient of the Analytically Unknown Value Function

MALKHAZ SHASHIASHVILI

Section 1. The Basic Idea of the Research Project

It is well known that vast majority of the real-world optimization problems cannot be solved analytically in closed form since they are highly nonlinear by their intrinsic nature.

Denote $V(x)$ the real−valued **Value Function** of the optimization problem, where we minimize corresponding **Objective Function** $F(x, z)$ over certain parametric set Z and the argument x belongs to some domain D of the n-dimensional Euclidean space R^n.

Numerical methods suggest to approximate analytically unknown **Value Function** $V(x)$ on a dense discrete subset G of grid points by the function $V(x, h)$, where x belongs to the discrete grid set G and the parameter h shows the "denseness" of G with respect to the domain D and is defined as the smallest positive number satisfying the following condition: for arbitrary point y belonging to domain D one can find a point x belonging to the grid set G such that $|y - x| < h$.

Many real-world problems also require the approximate computation of the **Gradient** $grad\, V(x)$, that is the vector of all partial derivatives of the function $V(x)$. In general it turns out to be very difficult problem to construct sophisticated algorithm to approximate $grad\, V(x)$, as the corresponding difference quotients start wild oscillations when the parameter h tends to 0 and one finds out soon that the latter quotients converge nowhere in the limit.

Our basic observation: The **Value function** $V(x)$ of the optimization problem is often convex (or semi convex) in multidimensional argument x (for example, in engineering thermodynamics it is **the Convex Envelope of the Gibbs free energy function**). Therefore we should use the advantage of **Convexity** to construct convergent numerical approximations to $grad\,V(x)$.

Our basic idea: Assume that $V(x)$ is a convex function. Replace the approximation $V(x,h)$ by some convex approximation $C(x,h)$ in a hope that the latter one **will better imitate** the shape of the unknown convex function $V(x)$ and hence the **gradient** *grad* $C(x,h)$ can be announced as the reasonable approximation to the unknown $grad\,V(x)$!

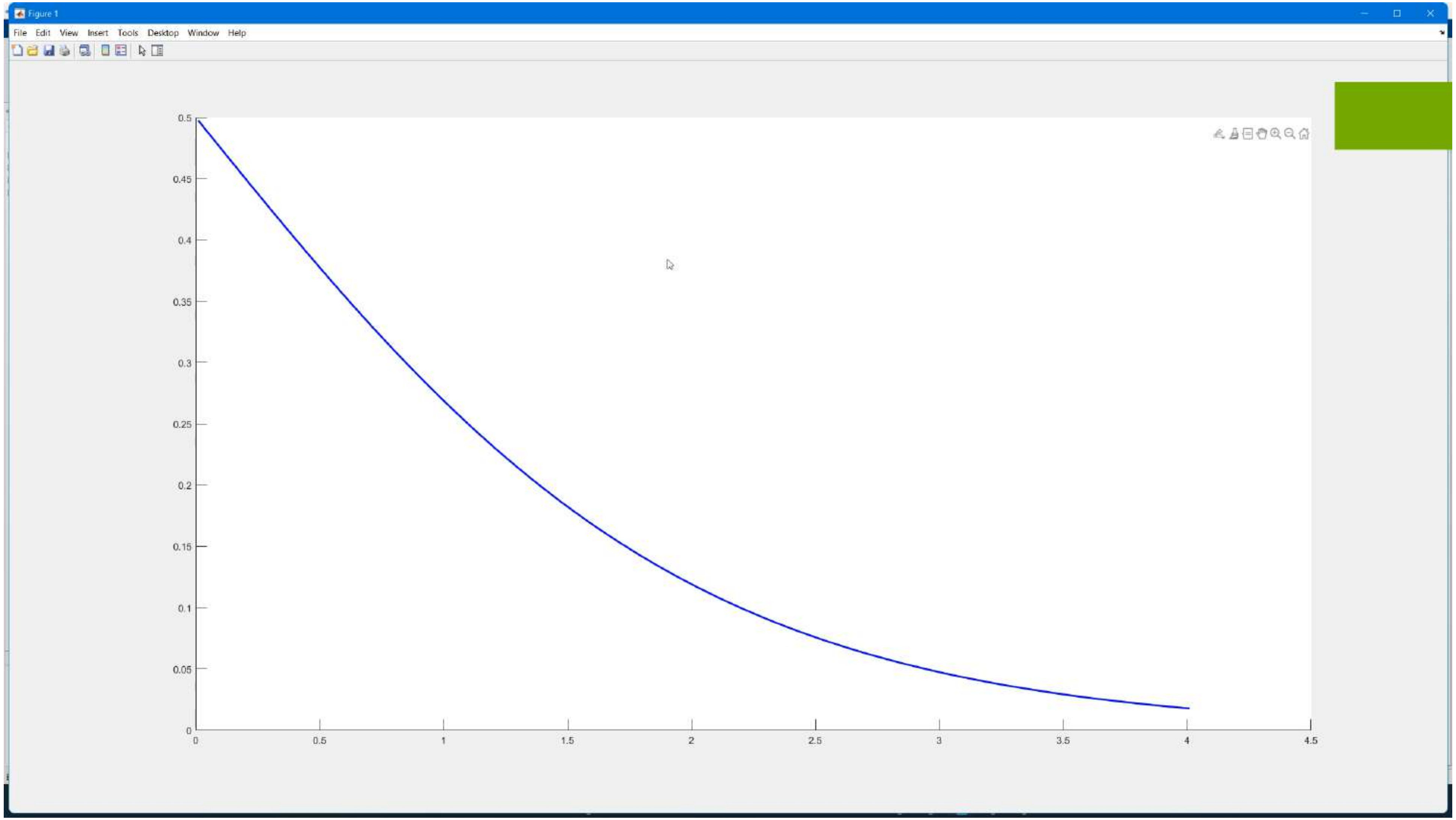
The clever choice of convex approximation consists in constructing the so called **Discrete Convex Envelope** denoted by $D\ conv\ V(x,h)$ of the function $V(x,h)$, which is defined on a domain D as the maximal convex function dominated by the function $V(x,h)$ **on a discrete set of grid points** G. The construction of **the discrete convex envelope** is carried out by several algorithms in computational geometry and most popular among them is **QHULL** (the quick hull algorithm for convex hulls), which finds the **convex hull** of arbitrary finite set of points in n- dimensional Euclidean space R^n and the **discrete convex envelope is obtained as a "lower part" of the corresponding convex hull** !

Our basic idea seems intuitively reasonable, but it needs rigorous mathematical justification. The latter justification has been given in our published paper

Shashiashvili K., Shashiashvili M. From the uniform approximation of a solution of the PDE to the L^2-approximation of the gradient of the solution. *J. Convex Anal.* **21** (2014), no. 1, 237-252,

where we have given rigorous mathematical justification of our intuitive arguments proving new type reverse Poincare inequalities for the **difference of two semi convex functions as well as for the difference of two convex envelopes of arbitrary continuous objective functions** not assuming even existence of first order partial derivatives of the **latter functions, see Proposition 3.2 and Theorem 3.3 therein**.

Section 2. Convex Envelope Animations



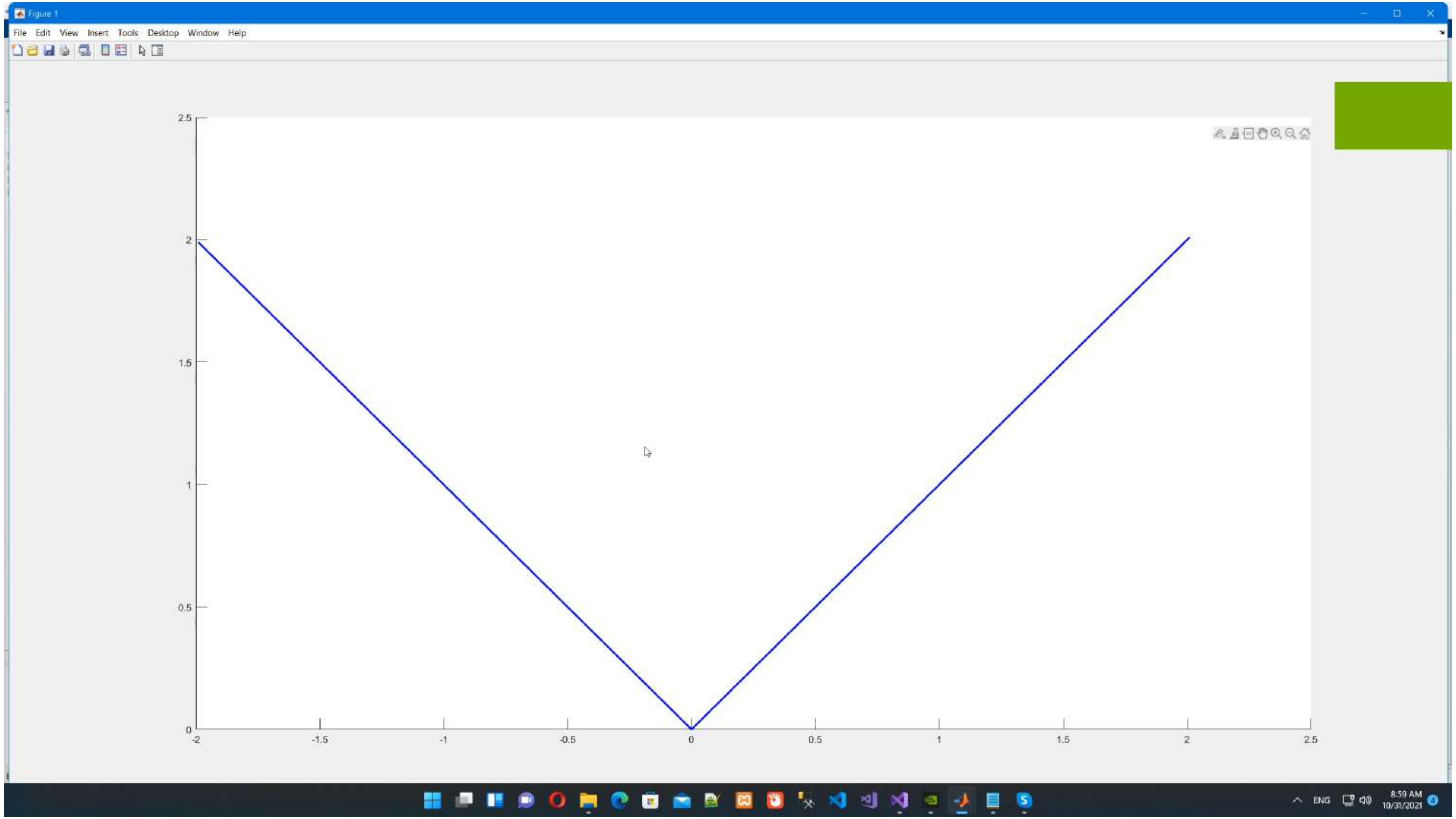


Figure 1
File Edit View Insert Tools Desktop Window Help
2.5
2
1.5
1
0.5
0
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
ENG
8:59 AM
10/31/2021

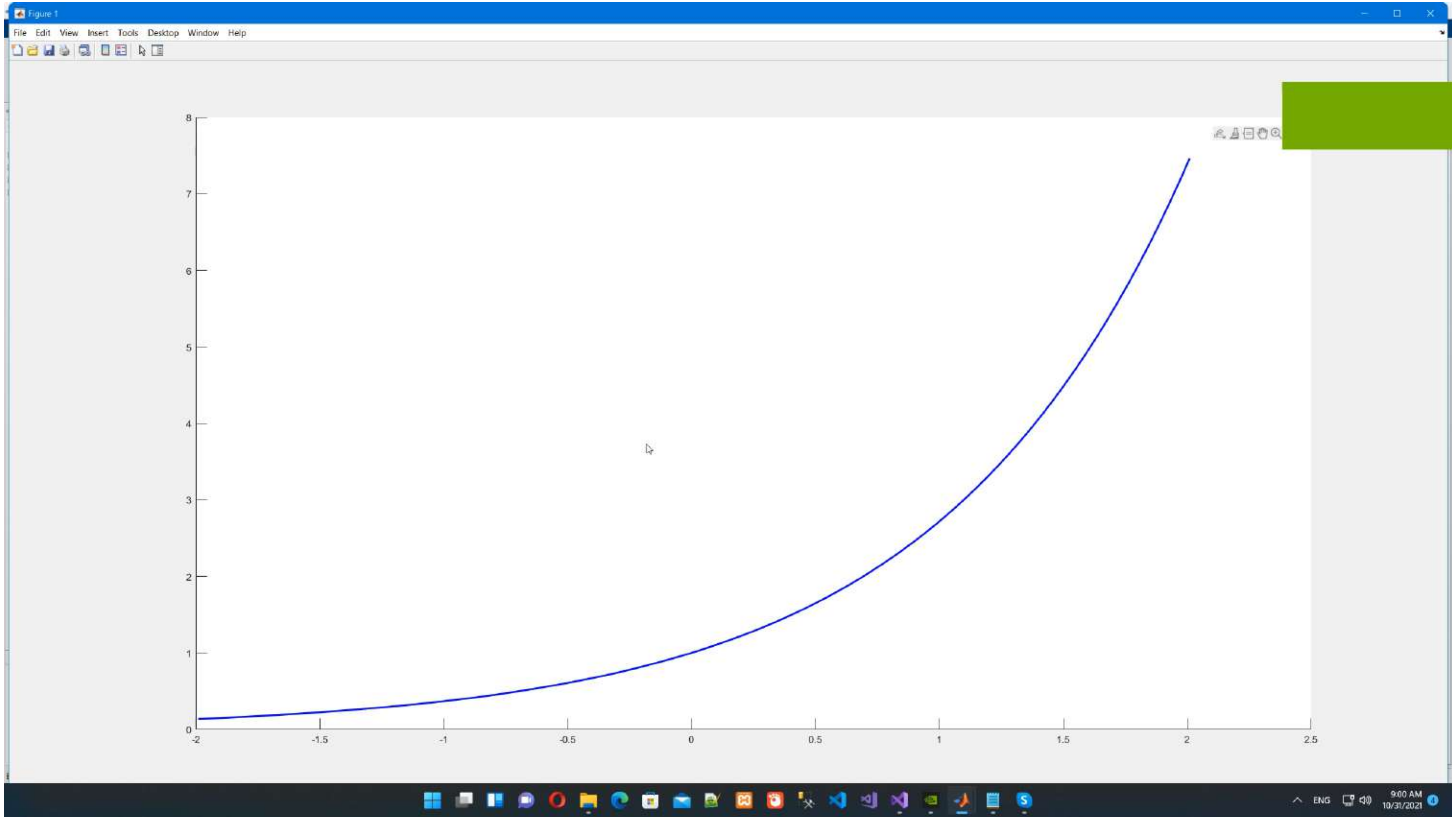


Figure 1
File Edit View Insert Tools Desktop Window Help
8
7
6
5
4
3
2
1
0
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5

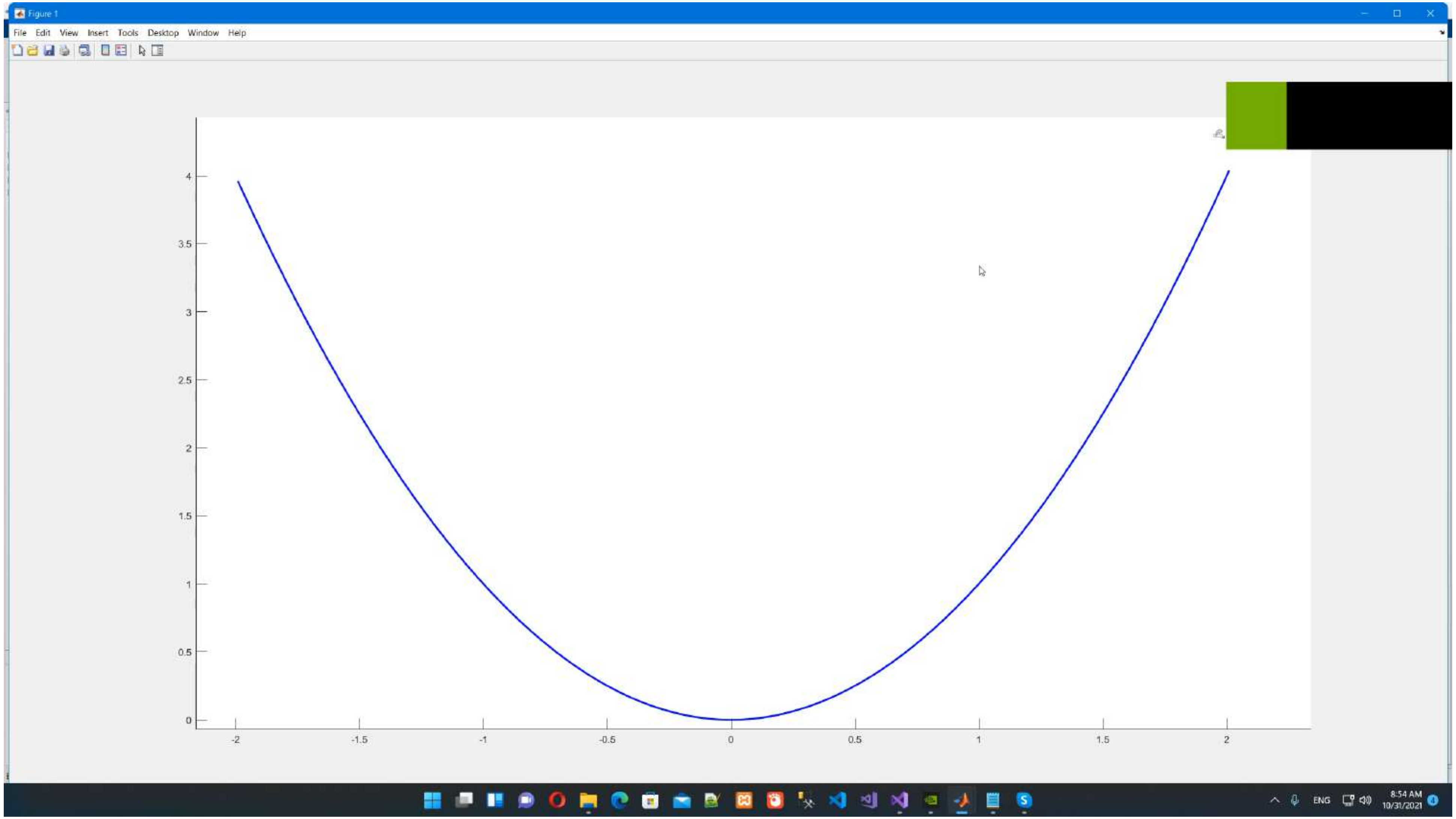


Figure 1
File Edit View Insert Tools Desktop Window Help
4
3.5
3
2.5
2
1.5
1
0.5
0
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
ENG
8:54 AM
10/31/2021

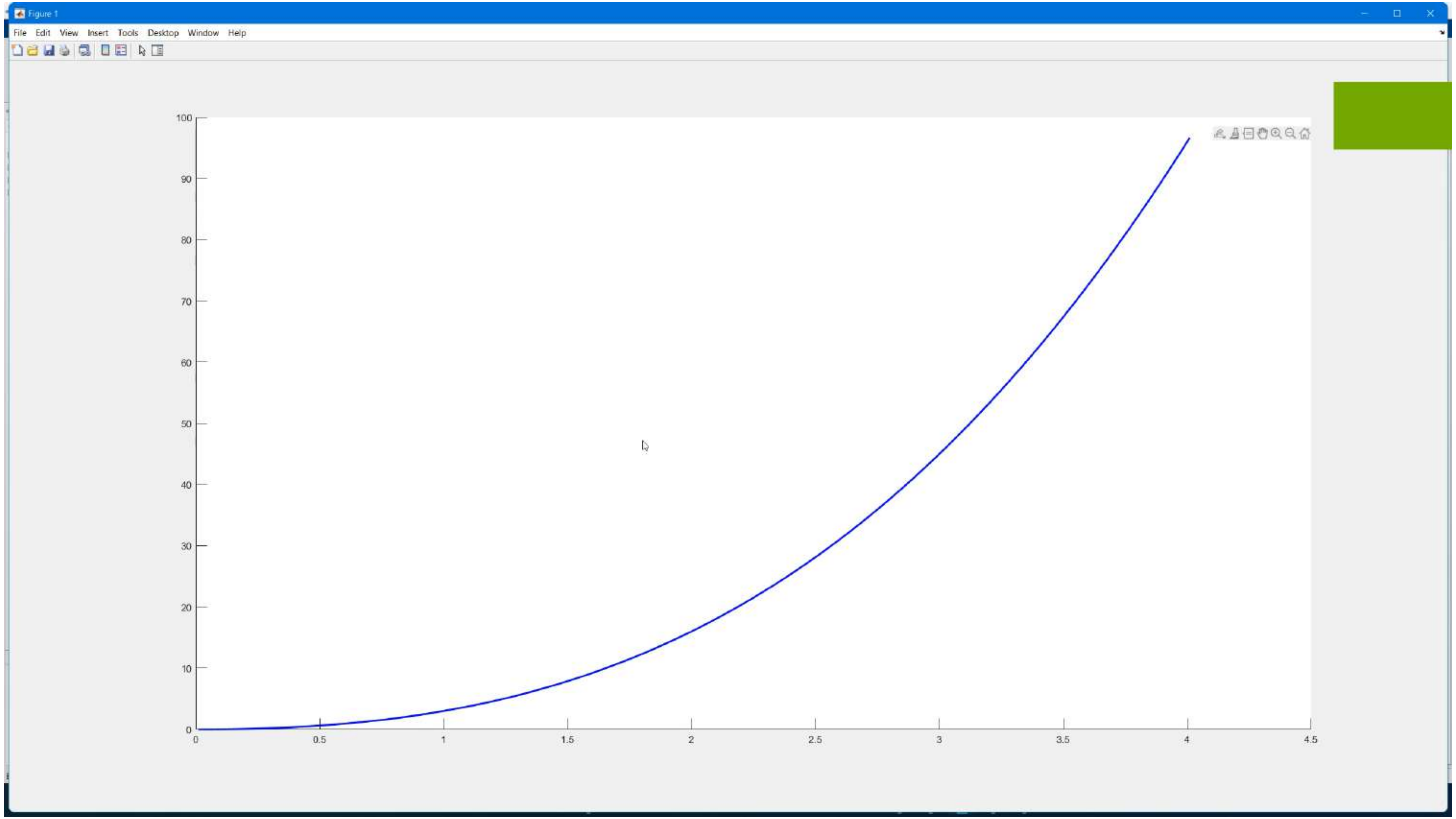


Figure 1
File Edit View Insert Tools Desktop Window Help
100
90
80
70
60
50
40
30
20
10
0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5

Section 3. The L^2-Approximation of the Gradient of the Semiconvex Function through the Convex Envelope

Let $u: D \to R$ be analytically unknown viscosity solution of the nonlinear second order elliptic partial differential equation

$$F(x, u, grad\ u, Hess\ u) = 0 \tag{3.1}$$

in a bounded open convex subset D of R^n.

As pointed out in the introduction the solution of the equation (3.1) turns out to be semiconvex (or semiconcave) function if the latter equation is related to different kind of optimization problems.

Suppose the bounded viscosity solution u of equation (3.1) is semiconvex function and we are given its uniform continuous numerical approximation $u_\delta: D \to R$, where δ is a small parameter, which typically measures the mesh size. The objective consists in constructing interior L^2-approximation of the unknown Sobolev gradient $grad\ u$ based on the uniform approximation u_δ. Moreover, it is desirable to estimate the gradient's L^2-error through the $L^\infty(D)$-uniform error of approximation.

We shall see in this section that such a construction is possible and it uses two ingredients: the energy inequality (2.5) in Shashiashvili M. and Shashiashvili K. [9] and the notion of the convex envelope.

The convex envelope $conv(u)$ of a bounded continuous function u in D is defined as the supremum of all convex functions which are majorized by the function u

$$conv(u) = \sup\{v(x): v(x)\ convex\ in\ D, v(x) \le u(x)\ for\ all\ x \in D\}\,. \tag{3.2}$$

The mapping $u \to conv(u)$ possesses some nice properties which we prove below

Lemma 3.1. *The mapping $u \to conv(u)$ has Lipschitz property*

$$\|conv(u) - conv(v)\|_{L^\infty(D)} \le \|u - v\|_{L^\infty(D)} \tag{3.3}$$

If only u, v belong to $C(D) \cap L^\infty(D)$.

Proof. Denote

$$d = \|u - v\|_{L^\infty(D)},$$

Then we have $-d \le u(x) - v(x) \le d$, i.e. $v(x) - d \le u(x)$, $u(x) - d \le v(x)$.

Hence we have

$$conv(v) - d \le u, conv(u) - d \le v.$$

This means that the convex functions $conv(v) - d$ and $conv(\mathrm{u}) - d$ are majorized respectively by u, v .

By the definition of the convex envelope we obtain

$$conv(v) - d \le conv(u), \qquad conv(u) - d \le conv(v),$$

i.e. $-d \le conv\ u(x) - conv\ v(x) \le d$, thus we derive the inequality (3.3). □

Taking successively $u = 0$ and $v = 0$ in (3.3) we get

$$\begin{cases} \|conv(v)\|_{L^\infty(D)} \leq \|v\|_{L^\infty(D)}, \\ \|conv(u)\|_{L^\infty(D)} \leq \|u\|_{L^\infty(D)}. \end{cases} \tag{3.4}$$

Proposition 3.2. *On the space $C(D) \cap L^\infty(D)$ the mapping $u \to conv(u)$ possesses the following important property*

$$\int_D |grad\ conv(u) - grad\ conv(v)|^2 \cdot \frac{d_{\partial D}^2}{n} dx \leq 5meas\ D \cdot \|u - v\|_{L^\infty(D)} \big(\|u\|_{L^\infty(D)} + \|v\|_{L^\infty(D)}\big). \tag{3.5}$$

Proof. We have from the bound (3.4) that the convex functions $conv(u)$ and $conv(v)$ are bounded, thus we can apply the energy inequality (2.5) for the latter convex functions and get

$$\int_D |grad\ conv(u) - grad\ conv(v)|^2 \cdot \frac{d_{\partial D}^2}{n} dx$$
$$\leq 5meas\ D \cdot \|conv(u) - conv(v)\|_{L^\infty(D)} \big(\|conv(u)\|_{L^\infty(D)} + \|conv(v)\|_{L^\infty(D)}\big). \tag{3.6}$$

The assertion follows after application of Lemma 3.1 and the bound (3.4). □

Consider now the bounded viscosity solution u of the equation (3.1) which is assumed to be semiconvex with semiconvexity constant $c \geq 0$ and its uniform continuous numerical approximation u_δ, i.e.

$$\|u_\delta - u\|_{L^\infty(D)} \underset{\delta \longrightarrow 0}{\longrightarrow} 0. \tag{3.7}$$

Further consider the bounded continuous functions

$$u + c \cdot v_0 \quad \text{and} \quad u_\delta + c \cdot v_0 \tag{3.8}$$

and their convex envelopes

$$\mathrm{conv}(u + c \cdot v_0) \quad \text{and} \quad conv(u_\delta + c \cdot v_0), \tag{3.9}$$

where

$$v_0(x) = \frac{1}{2} \cdot |x|^2.$$

The next proposition is the main result of Section 3.

Theorem 3.3. *The following weighted L^2-estimate is valid for the unknown $grad\, u$ through the function* $grad(conv(u_\delta + c \cdot v_0) - c \cdot v_0)$

$$\int_D |grad(conv(u_\delta + c \cdot v_0) - c \cdot v_0) - grad\, u|^2 \cdot \frac{d_{\partial D}^2}{n} dx$$
$$\leq 5meas\, D\|u_\delta - u\|_{L^\infty(D)}\left(2\|u\|_{L^\infty(D)} + \|u_\delta - u\|_{L^\infty(D)} + 2c\|v_0\|_{L^\infty(D)}\right). \quad (3.10)$$

Proof. Let us apply Proposition 3.2 to the functions $u_\delta + c \cdot v_0$ and $u + c \cdot v_0$, we shall have

$$\int_D |grad\, conv(u_\delta + c \cdot v_0) - grad\, conv(u + c \cdot v_0)|^2 \cdot \frac{d_{\partial D}^2}{n} dx$$
$$\leq 5meas\, D\|u_\delta - u\|_{L^\infty(D)}\left(\|u_\delta + c \cdot v_0\|_{L^\infty(D)} + \|u + c \cdot v_0\|_{L^\infty(D)}\right). \quad (3.11)$$

By the semiconvexity criteria (2.3) in Shashiashvili M. and Shashiashvili K. [9] we have that the function $u + c \cdot v_0$ is convex and therefore coincides with its convex envelope $conv(u + c \cdot v_0)$, hence we get

$$grad\, conv(u + c \cdot v_0) = grad(u + c \cdot v_0) = grad\, u + grad(c \cdot v_0),$$

the rest is obvious. □

Thus the L^2-approximation problem of the unknown $grad\, u$ is reduced to the efficient numerical computation of convex envelope $conv(u_\delta + c \cdot v_0)$ and its gradient. We note here that if the solution of PDE (3.1) is convex the unknown $grad\, u$ is approximated by the $grad\, conv(u_\delta)$.

Section 4. Computation of the Gradient of the Solution of Monge-Ampere Partial Differential Equation in a Planar Domain

We discuss next the Monge-Ampere equation. The Monge-Ampere equation is a fully nonlinear elliptic PDE. Applications of the Monge-Ampere equation appear in the classical problem of prescribed Gauss curvature and in the problem of optimal mass transportation (with quadratic cost).

We shall present a simple (nine point stencil) finite difference method which performs well for smooth as well as for singular solutions. The Monge-Ampere PDE in a planar domain $D \subset R^2$ is the following

$$det\big(Hessian\ U(x)\big) = f(x), f(x) \geq 0,$$

or equivalently

$$\frac{\partial^2 u}{\partial x^2} \cdot \frac{\partial^2 u}{\partial y^2} - \left(\frac{\partial^2 u}{\partial x \partial y}\right)^2 = f \quad \text{with Dirichlet boundary conditions } u = g \text{ on } \partial D \tag{4.1}$$

and the additional convexity constraint

$$u(x, y) \quad \text{is convex in } D, \tag{4.2}$$

which is required for the equation to be elliptic. Without the convexity constraint this equation does not have a unique solution. For example, taking the boundary function $g = 0$, if u is a solution, then $-u$ is also a solution.

The numerical method involves simply discretizing the second derivatives using standard central differences on a uniform Cartesian grid. The result is

$$\left(D_{xx}^2 u_{ij}\right) \cdot \left(D_{yy}^2 u_{ij}\right) - \left(D_{xy}^2 u_{ij}\right)^2 = f_{ij}, \tag{4.3}$$

where

$$\begin{cases} D_{xx}^2 u_{ij} = \dfrac{u_{i+1,j} + u_{i-1,j} - 2u_{ij}}{h^2}, \\ D_{yy}^2 u_{ij} = \dfrac{u_{i,j+1} + u_{i,j-1} - 2u_{ij}}{h^2}, \\ D_{xy}^2 u_{ij} = \dfrac{u_{i+1,j+1} + u_{i,j-1} - u_{i-1,j+1} - u_{i-1,j-1}}{4h^2}. \end{cases} \tag{4.4}$$

Introduce the notation

$$a_1 = \frac{u_{i+1,j} + u_{i-1,j}}{2}, \qquad a_2 = \frac{u_{i,j+1} + u_{i,j-1}}{2}, \qquad a_3 = \frac{u_{i+1,j+1} + u_{i,j-1}}{2}, \qquad a_4 = \frac{u_{i-1,j+1} + u_{i-1,j-1}}{2} \tag{4.5}$$

and rewrite (4.3) as a quadratic equation for u_{ij}:

$$4\left(a_1 - u_{ij}\right)\left(a_2 - u_{ij}\right) - \frac{1}{4}(a_3 - a_4)^2 = h^4 f_{ij}. \tag{4.6}$$

Now solving for u_{ij} and selecting the smaller one (in order to select the locally convex solution), we obtain

$$u_{ij} = \frac{1}{2}(a_1 + a_2) - \frac{1}{2}\sqrt{(a_1 - a_2)^2 + \frac{1}{4}(a_3 - a_4)^2 + h^4 f_{ij}}. \tag{4.7}$$

We can now use Gauss-Seidel iteration to find the fixed point of (4.7).

The Dirichlet boundary conditions are enforced at boundary grid points. The convexity constraint (4.2) is not enforced (beyond the selection of the positive root in (4.7).

Next we consider two exact solutions for the Monge-Ampere PDE (4.1), (4.2) on the square $[0,1] \times [0,1]$.

Example 4.1.

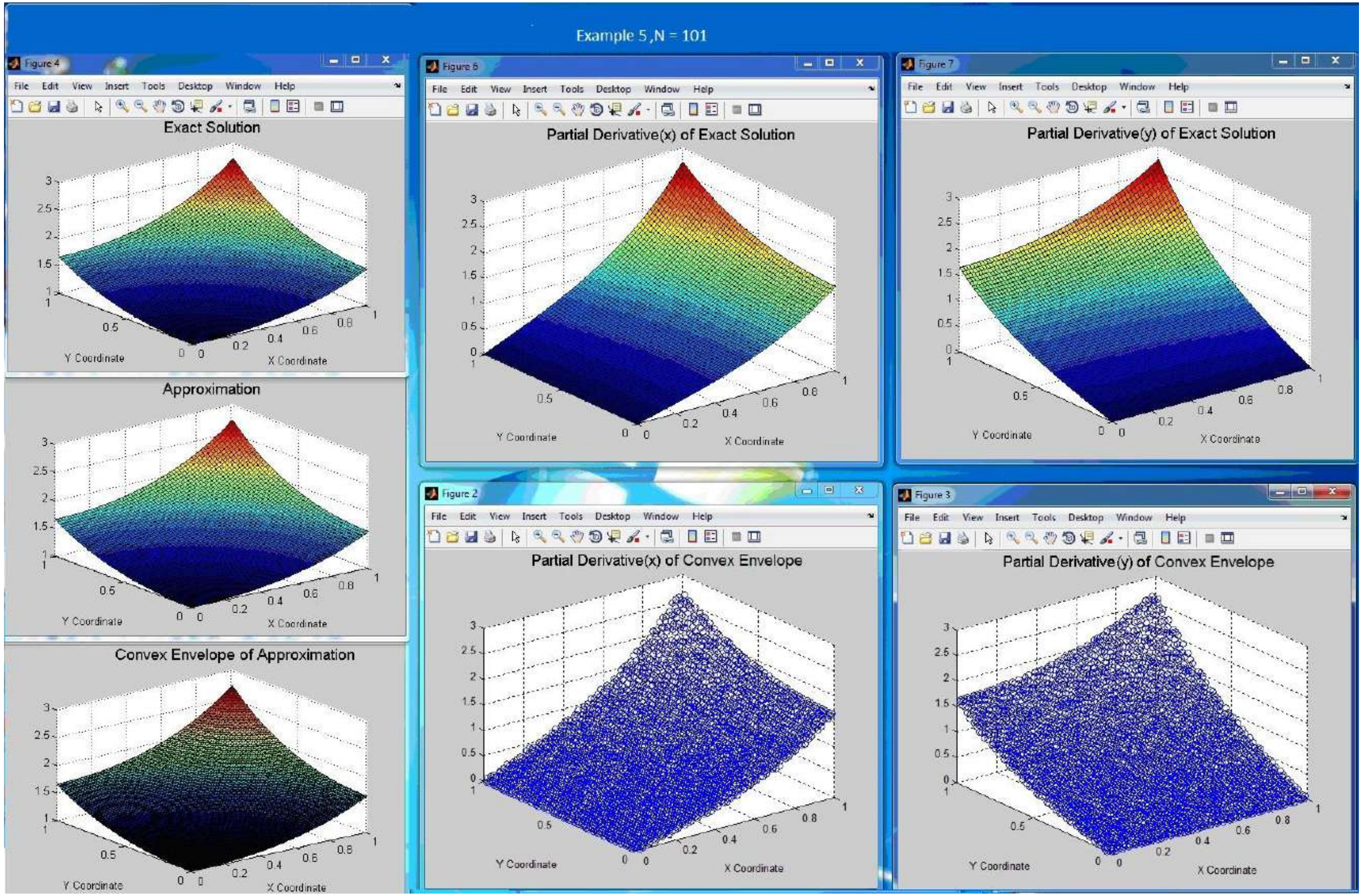
$$\begin{cases} u(x,y) = \exp\left(\dfrac{x^2+y^2}{2}\right), \\ f(x,y) = (1+x^2+y^2)\cdot\exp(x^2+y^2). \end{cases}$$

Example 4.2.

$$\begin{cases} u(x,y) = \dfrac{2\sqrt{2}}{3}(x^2+y^2)^{3/4}, \\ f(x,y) = \dfrac{1}{\sqrt{x^2+y^2}}. \end{cases}$$

In this example the function f blows up at the boundary point $(0,0)$.

We note that we use fast algorithm to accelerate computations in the finite difference method (4.3)-(4.7).



Example 5 ,N = 101
Figure 4
File Edit View Insert Tools Desktop Window Help
Exact Solution
Y Coordinate
X Coordinate
Approximation
Y Coordinate
X Coordinate
Convex Envelope of Approximation
Y Coordinate
X Coordinate
Figure 6
File Edit View Insert Tools Desktop Window Help
Partial Derivative(x) of Exact Solution
Y Coordinate
X Coordinate
Figure 7
File Edit View Insert Tools Desktop Window Help
Partial Derivative(y) of Exact Solution
Y Coordinate
X Coordinate
Figure 2
File Edit View Insert Tools Desktop Window Help
Partial Derivative(x) of Convex Envelope
Y Coordinate
X Coordinate
Figure 3
File Edit View Insert Tools Desktop Window Help
Partial Derivative(y) of Convex Envelope
Y Coordinate
X Coordinate

Figure 1

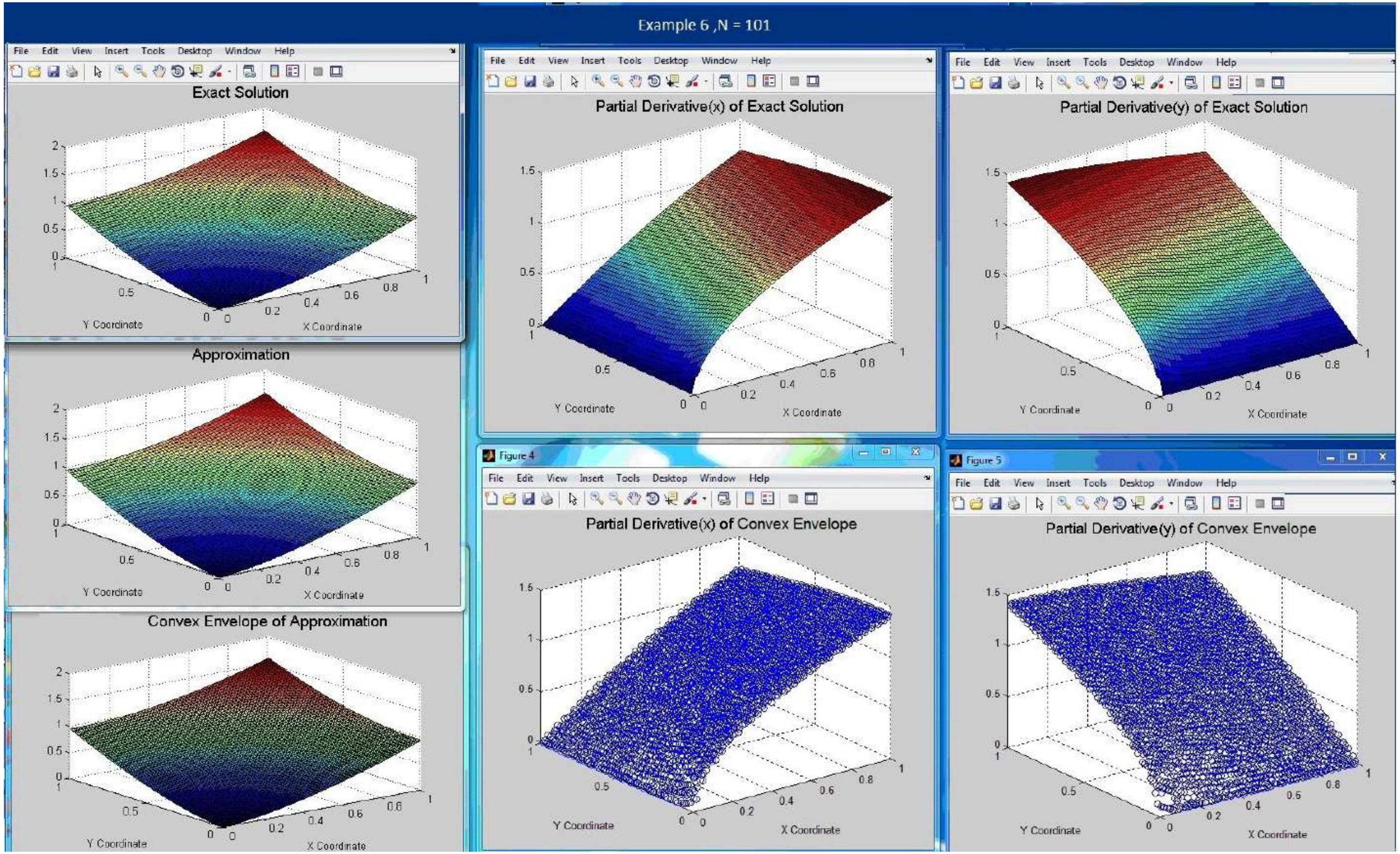


Figure 2

The Monge-Ampere equations (the Examples 4.1 and 4.2) are considered on the square $[0,1] \times [0,1]$.

In the tables below for the different grid points we compute the number of iterations, the computation times, the errors of approximation of the exact solution and of the exact gradient.

Computation times and errors for the exact solution and its gradient for the Example 4.1 on an $N \times N$ grid:

#	Number of iterations	Computation times	Uniform error for the exact solution	Uniform error for the exact gradient	L^2-error for the exact gradient
21	1362	1 sec.	1.5×10^{-4}	0.1255	0.011
61	10840	10 sec.	1.8×10^{-5}	0.0441	0.0038
101	28764	60 sec.	6.7×10^{-6}	0.0267	0.0023
141	54802	300 sec.	3.4×10^{-6}	0.0192	0.0016

Computation times and errors for the exact solution and its gradient for the Example 4.2 on an grid:

#	Number of iterations	Computation times	Uniform error for the exact solution	Uniform error for the exact gradient	L^2-error for the exact gradient
21	1397	1 sec.	1.5×10^{-4}	0.1511	0.0077
61	11065	10 sec.	1×10^{-4}	0.0887	0.0027
101	29312	70 sec.	4.9×10^{-5}	0.0689	0.0016
141	55768	300 sec.	2.9×10^{-5}	0.0583	0.0011

We give the surface plots (for Examples 4.1 and 4.2) of the following functions:

a) the exact solution,

b) finite difference numerical approximation,

c) the convex envelope of the numerical approximation,

d) partial derivative w.r. to x of the exact solution,

e) partial derivative w.r. to y of the exact solution,

f) partial derivative w.r. to x of the convex envelope,

g) Partial derivative w.r. to y of the convex envelope.

Section 5. Pricing and Hedging of American Options written on Multiple Assets

In this section we study the multidimensional parabolic obstacle problem and its relation to the pricing and hedging of American options written on multiple assets. We shall consider the so called strong solutions of parabolic obstacle problem that have been studied, for example, in Friedman [3, Chapter 1]. Strong solutions have second order Sobolev (weak) derivatives so that the Partial Differential Equation (PDE) can be written pointwisely a.e., strong solutions should be preferable in financial applications because of their better regularity properties.

The above obstacle problem appears naturally in the valuation of American type Claims in financial market. The obstacle is the so called payoff function and the solution of the obstacle problem is the value function of the American option written on multiple assets. A good background study is given in the paper by Broadie and Detemple [1].

The semiconvexity is a natural property of a large class of value functions of the optimization problems (see, for instance, Cannarsa and Sinestrari [2]).

This convexity (semiconvexity) of the value function of the American option for arbitrary fixed time instant is the starting point of our new method of the construction of the nearly optimal discrete time delta hedging strategies for American options.

American option can be exercised by its holder (as an opposite to European option) at any time up to and including expiry. This makes their pricing mathematically challenging and few closed form solutions have been found. American options are important because they are very widely traded. At least as important as the pricing of American options are the hedging issues that are crucial for the writer of the option.

In this section we study the parabolic obstacle problem in the strong sense. More precisely, we seek a solution $u(x,t)$, which belongs to the parabolic Sobolev space (see, for example, Krylov [6, Chapter 2]) and satisfies a system of inequalities

$$\begin{cases} Lu(x,t) \le 0, \qquad u(x,t) \ge g(x), \\ Lu(x,t) \cdot \big(u(x,t) - g(x)\big) = 0 \end{cases} \tag{5.1}$$

$(dx \times dt)$ with terminal condition

$$u(x,T) = g(x), \tag{5.2}$$

where $g(x)$, $x \in R^n$ is a given non-negative continuous function representing an obstacle and Lu is the second order linear parabolic differential operator

$$Lu(x,t) = \sum_{i,j=1}^{n} a_{ij}(x,t) \cdot \frac{\partial^2 u(x,t)}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x,t) \cdot \frac{\partial u(x,t)}{\partial x_i} - r(t) \cdot u(x,t) + \frac{\partial u(x,t)}{\partial t}, \tag{5.3}$$

when the obstacle $g(x)$ is non-smooth there are not many known techniques to be used in the study of the obstacle problem. Our objective is to develop some new results for the nonsmooth case, with focus on applications to American type options written on multiple assets, which is an active research area at present in mathematical finance.

We will consider the pricing and hedging of multidimensional American options in a financial market driven by a general multidimensional Ito diffusion. The American option is a financial contract, assuming a time horizon of $T > 0$ and a market consisting of n assets $S(t) = \left(S_1(t), \dots, S_n(t)\right)$ giving a payoff at time t equal to $\Psi\left(S_1(t), \dots, S_n(t)\right)$ where $\Psi(x)$ is a non-negative continuous function from R_+^n to R_+ defining the contract. The American option corresponding to this claim gives the owner of the option the right (but not the obligation) to exercise the option at any time τ, $0 \leq \tau \leq T$. At the exercise time τ, the owner of the option receives an amount equal to $\Psi\left(S(\tau)\right)$. We suppose the existence of a positive and continuous instantaneous interest rate $r(t)$ and also of the dividend rates $d_i(t)$ of the assets $S_i(t)$, $i = 1, \dots, n$.

We assume that there exists a risk-neutral martingale measure Q, such that with respect to Q the logarithms of the prices $X(t) = \left(\ln\left(S_1(t)\right), \dots, \ln\left(S_n(t)\right)\right)$ solve a system of stochastic differential equations

$$dX(t) = b(X(t), t) \cdot dt + \sigma(X(t), t) \cdot dW(t), \qquad X(0) = x, \qquad 0 \le t \le T, \tag{5.4}$$

where

$$b_i(x, t) = r(t) - d_i(t) - \frac{1}{2} \sum_{k=1}^{n} \sigma_{ik}^2, \qquad i = 1, \dots, n, \tag{5.5}$$

Here $W(t) = \left(W_1(t), \dots, W_n(t)\right)$ is a standard n- dimensional Brownian motion with respect to the filtration $(\mathfrak{J}_t)_{0 \le t \le T}$ defined on a probability space $(\Omega, \mathfrak{J}, Q)$, $\sigma(x, t) = \left(\sigma_{ij}(x, t)\right)_{i,j=1,\dots,n}$, where

$$a_{ij}(x, t) = \frac{1}{2} \sum_{k=1}^{n} \sigma_{ik}(x, t) \cdot \sigma_{jk}(x, t)\,. \tag{5.6}$$

We will assume that the operator Lu is uniformly parabolic in the sense that there exists $\lambda > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x,t) \cdot \xi_i \cdot \xi_j \geq \lambda \cdot |\xi|^2, \qquad \text{whenever} \quad (x,t) \in R^n \times [0,T] \quad \text{and} \quad \xi \in R^n. \tag{5.7}$$

We will assume also that the functions $b(x,t)$ and $\sigma(x,t)$ are bounded and Lipschitz continuous, that is, there exists a constant $c > 0$ such that for all $(x, \tilde{x}) \in R^n$ and $(s,t) \in [0,T]$ we have

$$\|\sigma(x,t) - \sigma(\tilde{x},s)\| + \|b(x,t) - b(\tilde{x},s)\| \leq c \cdot (\|x - \tilde{x}\| + |t - s|). \tag{5.8}$$

We will impose the basic assumption on the payoff function:

$$\Psi(x), \qquad x \in R_+^n \quad \text{is a nonnegative Lipschitz continuous convex function.} \tag{5.9}$$

Denote $V(x,t)$, $x \in R_+^n$, $0 \leq t \leq T$, the value function of the American option at time t, if the underlying assets are trading at $\big(S_1(t), \dots, S_n(t)\big) = (x_1, \dots, x_n)$. Then it is well known that

$$u(x,t) = V(\exp(x), t), \qquad x \in R^n, \qquad 0 \leq t \leq T, \tag{5.10}$$

is a unique solution of the parabolic obstacle problem (5.1), (5.2) with the obstacle function

$$g(x) = \Psi\big(\exp(x)\big), \qquad x \in R^n, \tag{5.11}$$

The convexity (semiconvexity) of the value function $V(x,t)$ of the American option for arbitrary fixed time instant t is the crucial point for our new device of the construction of the nearly optimal discrete time delta hedging strategies for American options written on multiple assets.

Indeed recently in the paper by Shashiashvili M. and Shashiashvili K. [9], we have developed a novel devise of numerical computation of the gradient of the analytically unknown function provided that the latter function is convex (or semiconvex) and we have already constructed its some uniform approximation. It is based on a new weighted inequality in Mathematical Analysis found by us (called otherwise the reverse Poincare inequality) for the difference of two semiconvex functions.

In this project we investigate the discrete time hedging problem for the American option written on the multiple underlying assets $S(t) = \big(S_1(t), \dots, S_n(t)\big)$, $0 \leq t \leq T$ and having a nonnegative convex payoff function $\Psi(x)$, $x \in R_+^n$.

It is a classical mathematical result at present (see, for example, Karatzas and Shreve [5, Chapter 2]) that for the perfect hedging in continuous time the writer of the option should construct the so called delta-hedging portfolio, which means that at an arbitrary time instant the should hold $grad\, V(S(t),t)$ units of the underlying assets, where $V(x,t)$ denotes the value function of the American option and $grad\, V(x,t)$ is a vector of its partial derivatives with respect to the components of its multidimensional space argument $x, x \in R_+^n$.

But the perfect hedging in continuous time requires the continuous rebalancing of the writer's portfolio in the underlying assets and the money market account, which is impossible in practice. In reality, the writer trades only at some discrete instants of time at which he rebalances his portfolio. Moreover, the delta-hedging requires the knowledge of the gradient $grad\, V(x,t)$ of the value function $V(x,t)$, but the explicit form neither of the value function, nor of its partial derivatives is known even in the simplest Black-Sholes model for American put option with finite horizon $T > 0$.

Several approximation methods were devised in order to compute the value function of the American option. In particular, finite difference methods were developed in Wilmott, Dewynne, and Howison [10], and Jaillet, Lamberton, and Lapeyre [4]. We assume here that we have already been given some continuous in the argument x uniform approximation $V_h(x,t)$ to the unknown value function $V(x,t)$ of the American option at the equidistant rebalancing times $t_k = k \cdot \delta$, $\delta = \frac{T}{N}$, $k = 0,1,\dots,N$ (for example, the Bermudan option value function approximation), where h is a certain small parameter indicating the error of approximation. In particular, we assume that the following bound is valid uniformly in k, $k = 0,1,\dots,N$,

$$\sup_{x\in R_+^n}|V_h(x,t_k) - V(x,t_k)| \leq c \cdot h, k = 0,1,\dots,N, \tag{5.12}$$

Here c is some positive constant depending on the parameters of our model and the payoff function $\Psi(x)$ and we naturally assume that

$$V_h(x,T) = \Psi(x), \qquad x \in R_+^n. \tag{5.13}$$

Our hedging method consists in the following: for each function $V_h(x, t_k)$, $k = 1,2, \dots, N$, consider first its convex envelope $conv\, V_h(x, t_k)$, $k = 1,2, \dots, N$, which is the maximal convex function dominated by the given function $V_h(x, t_k)$ and then its gradient $grad\ conv\, V_h(x, t_k)$, $k = 1,2, \dots, N$. Now the discrete time hedge $D_{\delta,h}(t)$, $0 \le t \le T$ can be defined in the following manner

$$D_{\delta,h}(t) = grad\ conv\, V_h(S(t_k), t_k) \ \text{ if } \ t_k \le t < t_{k+1}, \qquad k = 1, \dots, (N-1). \tag{5.14}$$

Our basic idea is to use the latter discrete time hedge as a reasonable approximation to the unknown continuous time optimal delta-hedge

$$D(t) = grad\, V(S(t), t), \qquad 0 \le t \le T. \tag{5.15}$$

Denote $\Pi_{\delta,h}(t)$, $0 \le t \le T$ the value process of the discrete time hedging portfolio and $\Pi(t)$, $0 \le t \le T$, respectively, the value process of continuous time optimal delta-hedging portfolio. Then the error due to our discrete time hedge is equal to

$$E^Q \sup_{0 \le t \le T} \left| \Pi_{\delta,h}(t) - \Pi(t) \right|. \tag{5.16}$$

One of the objectives of this research project consists in estimating the latter error for American options written on multiple assets and proving that it converges to zero when discretization parameters δ and h tend to zero. We should note here that this program has been successfully carried out in one dimensional case for Black-Sholes model in Shashiashvili and Hussain [7]. The estimation of the error (5.16) for multi asset American option problem will heavily rely on the weighted reverse Poincare inequalities in R^n and R_+^n and therefore proving such kind of inequalities is one of the objectives of this research project. We formulate the latter inequality in R^n.

Let $U(x)$ and $V(x)$ be two semiconvex functions in R^n with the semiconvexity constants c_U and c_V, respectively (see Cannarsa and Sinestrari [2, Chapter 1, Definition 1.1.1]) and $H(x)$ be a nonnegative twice continuously differentiable weight function. Then the following weighted reverse Poincare inequality should be valid (under certain conditions on $U(x)$, $V(x)$ and $H(x)$) for the difference $U(x) - V(x)$

$$\int_{R^n} |grad\, U(x) - grad\, V(x)|^2 \cdot H(x) dx$$

$$\leq c \|U - V\|_{L^\infty(R^n)} \int_{R^n} |U(x) + V(x) + max(c_U, c_V) \cdot |x|^2| \cdot |\Delta H(x)| dx\,, \qquad (5.17)$$

where Δ denotes the Laplace operator and c is the absolute constant.

References

1. Broadie M., Detemple J. The valuation of American options on multiple assets. *Math. Finance* **7** (1997), no. 3, 241-286.

2. Cannarsa P., Sinestrari C. *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*. Progress in Nonlinear Differential Equations and their Applications, 58. *Birkhäuser Boston, Inc., Boston, MA,* 2004.

3. Friedman A. Variational Principles and Free-Boundary Problems. Second edition. *Robert E. Krieger Publishing Co., Inc., Malabar, FL,* 1988.

4. Jaillet P., Lamberton D., Lapeyre B. Variational inequalities and the pricing of American options. *Acta Appl. Math.* **21** (1990), no. 3, 263-289.

5. Karatzas I., Shreve S. E. *Methods of Mathematical Finance*. Applications of Mathematics (New York), 39. *Springer-Verlag, New York,* 1998.

6. Krylov N. V. *Lectures on elliptic and parabolic equations in Sobolev spaces*. Graduate Studies in Mathematics, 96. American Mathematical Society, Providence, RI, 2008.

7. Hussain S., Shashiashvili M. Discrete time hedging of the American option. *Math. Finance* **20** (2010), no. 4, 647-670.

8. Rogava J., Shashiashvili K., Shashiashvili M. Computational convex analysis and its applications to numerical solution of some nonlinear partial differential equations. *Appl. Math. Inform. Mech.* **19** (2014), no. 2, 62-82.

9. Shashiashvili K., Shashiashvili M. From the uniform approximation of a solution of the PDE to the L^2-approximation of the gradient of the solution. *J. Convex Anal.* **21** (2014), no. 1, 237-252.

10. Wilmott P., Dewynne J., Howison S. *Option Pricing: Mathematical Models and Computations*. Oxford Financial Press, Oxford, UK, 1993.

Thank you very much for your attention!

The Adomian series representation of some quadratic BSDEs

R. Tevzadze

Abstract. The representation of the solution of some Backward Stochastic Differential Equation as an infinite series is obtained. Some exactly solvable examples are considered.

2020 Mathematics Subject Classification. 90A09, 60H30, 90C39

Keywords: Stochastic exponential, martingale, Adomian series, Brownian Motion.

1 Introduction

In a number of papers[1,2] Adomian develops a numerical technique using special kinds of polynomials for solving non-linear functional equations. However, Adomian and his collaborators did not develop widely the problem of convergence.

In this article we will study by Adomian technique some kind of quadratic backward martingale equation and prove the convergence of the series. For example we tackle an equation of the form

$$\mathcal{E}_T(m)\mathcal{E}_T^{\alpha}(m^{\perp}) = c\exp\{\eta\} \tag{1}$$

w.r.t. stochastic integrals $m = \int f_s dW_s,\ m^{\perp} = \int g_s dW_s^{\perp}$ and real number c, where $(W, W^{\perp})$ is 2-dimension Brownian Motion and η is a random variable.

Equations of such type are arising in mathematical finance and they are used to characterize optimal martingale measures (see, Biaginiat at al (2000), Mania and Tevzadze (2000), (2003),(2006)). Note that equation (1) can be applied also to the financial market models with infinitely many assets (see M. De Donno at al (2003)). In Biagini at al (2000) an exponential equation of the form

$$\frac{\mathcal{E}_T(m)}{\mathcal{E}_T(m^\perp)} = ce^{\int_0^T \lambda_s^2 ds}$$

was considered (which corresponds to the case $\alpha = -1$).

Our goal is to show the solvability of the equation (1) using the Adomian method proving the convergence of series. On the one hand, a simpler proof of solvability is obtained. On the other hand, it allows to obtain the approximation of the solution. It is possible to find a solution in the form of series, if we define a sequence of martingales w.r.t. the measure $\mathcal{E}_T(\sum_i^n m_i + \sum_i^n m_i^\perp)\cdot P$ from equations $c'\mathcal{E}_T(m'_{n+1} + m'^\perp_{n+1}) = \mathcal{E}_T^2(m'^\perp_n)$, where $m'_{n+1} = m_{n+1} - \langle m_{n+1}, \sum_i^n, m_i\rangle$, $m'^\perp_{n+1} = m^\perp_{n+1} - \langle m^\perp_{n+1}, \sum_i^n m_i^\perp\rangle$, and then we write down the solution

$$m = \sum_k^\infty m_k,\ m^\perp = \sum_k^\infty m_k^\perp$$

provided the series are convergent. The proof of the convergence is greatly simplified if we present equation as a BSDE in the space of BMO-martingales and use the properties of the BMO-norm. The result is resumed in Theorem 1.

Finally we provide some examples, exactly solvable by Adomian series and also example non-solvable at all.

2 The main result

Let $(\Omega, \mathcal{F}, P)$ be a probability space with filtration $\mathbf{F} = (\mathcal{F}_t, t \in [0, T])$. We assume that all local martingales with respect to $\mathbf{F}$ are continuous. Here T is a fixed time horizon and $\mathcal{F} = \mathcal{F}_T$.

Let $\mathcal{M}$ be a stable subspace of the space of square integrable martingales H^2. Then its ordinary orthogonal $\mathcal{M}^\perp$ is a stable subspace and any element of $\mathcal{M}$ is strongly orthogonal to any element of $\mathcal{M}^\perp$ (see, e.g. [5], [6]).

We consider the following exponential equation

$$\mathcal{E}_T(m)\mathcal{E}_T^\alpha(m^\perp) = c\exp\{\eta\}, \tag{2}$$

where η is a given F_T-measurable random variable and α is a given real number. A solution of equation (2) is a triple $(c, m, m^{\perp})$, where c is strictly positive constant, $m \in \mathcal{M}$ and $m^{\perp} \in \mathcal{M}^{\perp}$. Here $\mathcal{E}(X)$ is the Doleans-Dade exponential of X.

It is evident that if $\alpha = 1$ then equation (2) admits an "explicit" solution. E.g., if $\alpha = 1$ and η is bounded, then using the unique decomposition of the martingale $E(\exp\{\eta\}/F_t)$

$$E(\exp\{\eta\}/F_t) = E\exp\{\eta\} + m_t(\eta) + m_t^{\perp}(\eta), \quad m(\eta) \in \mathcal{M}, \quad m^{\perp}(\eta) \in \mathcal{M}^{\perp}, \tag{3}$$

it is easy to verify that the triple $c = \frac{1}{E\exp\{\eta\}}$,

$$m_t = \int_0^t \frac{1}{E(\exp\{\eta\}/F_s)} dm_s(\eta), \quad m_t^{\perp} = \int_0^t \frac{1}{E(\exp\{\eta\}/F_s)} dm_s^{\perp}(\eta)$$

satisfies equation (2).

Our aim is to prove the existence of a unique solution of equation (2) for arbitrary $\alpha \neq 0$ and η of a general structure, assuming that it satisfies the following boundedness condition:

B) η is an F_T-measurable random variable of the form

$$\eta = \bar{\eta} + \gamma A_T, \tag{4}$$

where $\bar{\eta} \in L^{\infty}$, γ is a constant and $A = (A_t, t \in [0, T])$ is a continuous F-adapted process of finite variation such that

$$E(var_T(A) - var_{\tau}(A)/F_{\tau}) \leq C$$

for all stopping times τ for a constant $C > 0$.

One can show that equation (2) is equivalent to the following semimartingale backward equation with the square generator

$$Y_t = Y_0 - \frac{\gamma}{2}A_t - \langle L \rangle_t - \frac{1}{\alpha}\langle L^{\perp} \rangle_t + L_t + L_t^{\perp}, \quad Y_T = \frac{1}{2}\bar{\eta}. \tag{5}$$

We use also the equivalent equation of the form

$$L_T + L_T^{\perp} = c + \langle L \rangle_T + \frac{1}{\alpha}\langle L^{\perp} \rangle_T + \frac{\gamma}{2}A_T.$$

w.r.t. $(c, L, L^{\perp})$.

We use notations $|M|_{\text{BMO}} = \inf\{C : E^{\frac{1}{2}}(\langle M\rangle_T - \langle M\rangle_\tau|\mathcal{F}_\tau) \le C\}$ for BMO-norms of martingales, $|A|_\omega = \inf\{C : E(var_t^T(A)|\mathcal{F}_t) \le C\}$ for norms of finite variation processes and $A\cdot M$ for stochastic integrals.

Let us consider the system of semimartingale backward equations

$$Y_t^{(0)} = Y_0^{(0)} - \frac{\gamma}{2}A_t + L_t^{(0)} + L_t^{(0)\perp}, \quad Y_T^{(0)} = \frac{1}{2}\bar\eta,$$

$$Y_t^{(n+1)} = Y_0^{(n+1)}$$

$$-\sum_{k=0}^{n}\langle L^{(k)}, L^{(n-k)}\rangle_t - \frac{1}{\alpha}\sum_{k=0}^{n}\langle L^{(k)\perp}, L^{(n-k)\perp}\rangle_t + L_t^{(n+1)} + L_t^{(n+1)\perp},$$

$$Y_T^{(n+1)} = 0.$$

The sequence $Y_0^{(n)} = c^{(n)}, L^{(n)} + L^{\perp(n)}, n = 0, 1, 2, \cdots$ can be defined consequently by the equations

$$E(\eta|\mathcal{F}_t) + \frac{\gamma}{2}E(A_T|\mathcal{F}_t) = c^{(0)} + L_t^{(0)} + L_t^{\perp(0)},$$

$$\sum_{k=0}^{n}E(\langle L^{(k)}, L^{(n-k)}\rangle_T|\mathcal{F}_t) - \frac{1}{\alpha}\sum_{k=0}^{n}E(\langle L^{(k)\perp}, L^{(n-k)\perp}\rangle_T|\mathcal{F}_t)$$

$$= c^{(n+1)} + L_t^{(n+1)} + L_t^{\perp(n+1)}.$$

Remark. If $A_t = \int_0^t a(s, W_s, B_s)ds$, then the solution of (5) is of the form $Y_t = v(t, W_t, B_t)$, where $v(t, x, y)$ is decomposed as series $\sum_n v^n(t, x, y)$ satisfying the system of PDEs

$$(\partial_t + \frac{1}{2}\Delta)v^0(t, x, y) + a(t, x, y) = 0, \quad v^0(T, x, y) = 0,$$

$$(\partial_t + \frac{1}{2}\Delta)v^n(t, x, y)$$

$$+\frac{1}{2}\sum_{k=0}^{n-1}(v_x^k(t, x, y)v_x^{n-k-1}(t, x, y) + \alpha v_y^k(t, x, y)v_y^{n-k-1}(t, x, y)) = 0,$$

$$v^n(T, x, y) = 0,\ n \ge 1.$$

Lemma 1. *Let*

$$Y_t = Y_0 + A_t + m_t, \quad Y_T = \eta,$$

where m is a martingale, $\eta \in L_\infty$ and $|A|_\omega < \infty$. Then $m \in BMO$ and

$$|m|_{\text{BMO}} \le |\eta|_\infty + |A|_\omega. \tag{6}$$

In particular, if $|A|_\omega < \infty$ *then the martingale* $E(A_T|F_t)$ *belongs to the BMO space and*

$$|E(A_T|F_\cdot)|_{\text{BMO}} \le |A|_\omega.$$

Proof. By the Ito formula

$$Y_t^2 = 2\int_0^t Y_s dm_s + 2\int_0^t Y_s dA_s + \langle m\rangle_t.$$

Taking the difference $Y_\tau^2 - Y_T^2$ and conditional expectations we have that

$$Y_\tau^2 + E(\langle m\rangle_T - \langle m\rangle_\tau|F_\tau) = E(\eta^2|F_\tau) - 2E(\int_\tau^T Y_s dA_s|F_\tau) \le$$

$$\le |\eta|_\infty^2 + 2|Y|_\infty|A|_\omega. \tag{7}$$

$E(\int_\tau^T Y_s dm_s|F_\tau) = 0$, since $Y_t \le E(\eta + |A_T - A_t||\mathcal{F}_t)$ is bounded and m is a martingale. Since the right-hand side of (7) does not depend on τ from (7) we obtain

$$|Y|_\infty^2 + ||m||_{BMO}^2 \le |\eta|_\infty^2 + |Y|_\infty^2 + |A|_\omega^2.$$

Therefore

$$||m||_{BMO}^2 \le |\eta|_\infty^2 + |A|_\omega^2,$$

which implies inequality (6).

Lemma 2. *For the BMO norms of martingales* $L^{(n)} + L^{\perp(n)}$*, defined above, the following estimates are true*

$$|L^{(n)} + L^{\perp(n)}|_{\text{BMO}} \le a_n(1+|\beta|)^n|L^{(0)} + L^{\perp(0)}|_{\text{BMO}}^{n+1}, \tag{8}$$

where the coefficients a_n *are calculating recurrently from*

$$a_0 = 1, \quad a_{n+1} = \sum_{k=0}^n a_k a_{n-k}.$$

Proof. Using Lemma 1 it is easy to show that

$$|L^{(1)} + L^{\perp(1)}|_{\text{BMO}} \le a_1(1+|\beta|)|L^{(0)} + L^{\perp(0)}|_{\text{BMO}}^2,$$

$$|L^{(2)} + L^{\perp(2)}|_{\text{BMO}} \le a_2(1+|\beta|)^2|L^{(0)} + L^{\perp(0)}|_{\text{BMO}}^3.$$

Assume that inequality (8) is valid for any $k \leq n$ and let us show that

$$|L^{(n+1)} + L^{\perp(n+1)}|_{\text{BMO}} \leq a_{n+1}(1+|\beta|)^{n+1}|L^{(0)} + L^{\perp(0)}|^{n+2}_{\text{BMO}}. \tag{9}$$

Applying Lemma 1 for $Y_t^{(n+1)}$ and the Kunita-Watanabe inequality we have

$$|L^{(n+1)} + L^{\perp(n+1)}|_{\text{BMO}} \leq$$

$$\leq \operatorname*{ess\,sup}_{\tau} \sum_{k=0}^{n} E(var_\tau^T(\sum_k^n \langle L^{(k)}, L^{(n-k)}\rangle + \beta\langle L^{\perp(k)}, L^{\perp(n-k)}\rangle)|\mathcal{F}_\tau)$$

$$\leq \sum_{k=0}^{n} \operatorname*{ess\,sup}_{\tau} E^{\frac{1}{2}}(var_\tau^T\langle L^{(k)}\rangle|\mathcal{F}_\tau)E^{\frac{1}{2}}(var_\tau^T\langle L^{\perp(n-k)}\rangle|\mathcal{F}_\tau)$$

$$+|\beta|\sum_{k=0}^{n} \operatorname*{ess\,sup}_{\tau} E^{\frac{1}{2}}(var_\tau^T\langle L^{\perp(k)}\rangle|\mathcal{F}_\tau)E^{\frac{1}{2}}(var_\tau^T\langle L^{\perp(n-k)}\rangle|\mathcal{F}_\tau)$$

$$\leq \sum_{k}^{n} |L^{(k)}|_{\text{BMO}}|L^{(n-k)}|_{\text{BMO}} + |\beta||L^{\perp(k)}|_{\text{BMO}}|L^{\perp(n-k)}|_{\text{BMO}}$$

$$\leq (1+|\beta|)\sum_{k=0}^{n} |L^{(k)} + L^{\perp(k)}|_{\text{BMO}}|L^{(n-k)} + L^{\perp(n-k)}|_{\text{BMO}}. \tag{10}$$

Therefore, from (10), using inequalities (8) for any $k \leq n$, we obtain

$$|L^{(n+1)} + L^{\perp(n+1)}|_{\text{BMO}} \leq$$

$$\leq (1+|\beta|)\sum_{k=0}^{n} a_k(1+|\beta|)^k|L^{(0)}+L^{\perp(0)}|^{k+1}_{\text{BMO}}a_{n-k}(1+|\beta|)^{n-k}||L^{(n-k)}+L^{\perp(n-k)}|^{n-k+1}_{\text{BMO}}$$

$$\leq (1+|\beta|)^{n+1}|L^{(0)} + L^{\perp(0)}|^{n+2}_{\text{BMO}}\sum_{k=0}^{n} a_k a_{n-k} =$$

$$= a_{n+1}(1+|\beta|)^{n+1}|L^{(0)} + L^{\perp(0)}|^{n+2}_{\text{BMO}}$$

and the validity of inequality (8) follows by induction.

Theorem 1. *The series $\sum_{n\geq 0}(L^{(n)} + L^{\perp(n)})$ is convergent in BMO-space, if γ and $|\bar{\eta}|_\infty$ are small enough and the sum of series is a solution of the equation (5).*

Proof. Without loss of generality assume that $\eta = 0$. Using the lemma 2 we get

$$|L^{(n)} + L^{\perp(n)}|_{\text{BMO}} \le a_n(1+|\beta|)^n |L^{(0)} + L^{\perp(0)}|^{n+1}_{\text{BMO}} \le a_n(1+|\beta|)^n |\gamma A|^{n+1}_{\omega}.$$

By lemma 3 of appendix, since

$$\overline{\lim}_{n\to\infty} \sqrt[n]{a_n} = \overline{\lim}_{n\to\infty} \sqrt[n]{\frac{1}{2n+1} C^{2n+2}_{n+1}} = \overline{\lim}_{n\to\infty} \sqrt[n]{\frac{(2n)!}{n!n!}} = \overline{\lim}_{n\to\infty} \sqrt[n]{\frac{(2n)^{2n}}{n^{2n}}} = 4,$$

the series is convergent, when $\gamma < \frac{1}{4|A|_\omega(1+|\beta|)}$.

Remark. Since $\max(|L|_{\text{BMO}}, |L^\perp|_{\text{BMO}}) \le |L + L^\perp|_{\text{BMO}} \le |L|_{\text{BMO}} + |L^\perp|_{\text{BMO}}$ the convergence $\sum_{n\ge0}(L^{(n)} + L^{\perp(n)})$ implies convergence of $\sum_{n\ge0} L^{(n)}$ and $\sum_{n\ge0} L^{\perp(n)}$ and vice versa.

The existence of the solution for arbitrary bounded η is proven [8]. We can prove here little more general result

Proposition 1. *There exists solution of (2) for sufficiently small γ and arbitrary bounded $\bar\eta$.*

Proof. Let $\bar m + \bar m^\perp$ be solution of (2) for $\eta = \gamma A_T$ and sufficiently small γ. From the result of [8] there exists a solution of

$$\mathcal{E}_T(\tilde m)\mathcal{E}^\alpha_T(\tilde m^\perp) = c\exp\{\bar\eta\},$$

w.r.t

$$\bar P = \mathcal{E}_T(\bar m + \bar m^\perp)\cdot,\ \tilde m + \tilde m^\perp \in \mathcal{M}(F,\bar P) + \mathcal{M}^\perp(F,\bar P).P.$$

It is easy to verify that $m + m^\perp = \bar m + \bar m^\perp + \tilde m + \tilde m^\perp$ is a solution of (2) for $\eta = \bar\eta + \gamma A_T$.

The uniqueness of the solution was proved in [8].

Proposition 2. . *Let η be an $\mathcal{F}_T$ -measurable random variable. If there exists a triple $(c, m, m^\perp)$, where $c \in R_+, m \in BMO \cap \mathcal{M}, m^\perp \in BMO \cap \mathcal{M}^\perp$ satisfying equation (2) then such solution is unique.*

We now show that without finiteness of $|A|_\omega$ either the solution does not exists or the convergence of series is valid in a week sense.

Example 1. Let $\alpha = -1,\ \gamma = 2,\ \bar\eta = 0,\ A_t = \frac{1}{2}\int_0^t (W_s^2 + W_s^{2\perp})ds,\ \mathbf{F} = (\mathcal{F}_t^{W,W^\perp})$, where $W, W^\perp$ is 2-dimensional Brownian motion. Then (5) becomes

$$L_T + L_T^\perp = c + \langle L\rangle_T - \langle L^\perp\rangle_T + \frac{1}{2}\int_0^T (W_s^2 + W_s^{2\perp})ds.$$

We have

$$L_T^{(0)} + L_T^{(0)\perp} = c_0 + \int_0^T (T-s) W_s dW_s + \int_0^T (T-s) W_s^{\perp} dW_s^{\perp},$$

$$L_T^{n+1)} + L_T^{(n+1)\perp} = c_n + \sum_{k=0}^{n} \langle L^{(k)}, L^{(n-k)} \rangle_T - \sum_{k=0}^{n} \langle L^{(k)\perp}, L^{(n-k)\perp} \rangle_T,\ n \geq 0.$$

Let assume

$$L_T^{(n)} = \int_0^T (T-s)^{2n+1} \alpha_n W_t dW_s,$$
$$L_T^{(n)\perp} = \int_0^T (T-s)^{2n+1} \beta_n W_t^{\perp} dW_s^{\perp}.$$

Then $a_0 = 1,\ \ \beta_0 = 1$ and

$$L_T^{(n+1)} = c_n' + \sum_{k=0}^{n} \int_0^T (T-s)^{2n+2} \alpha_k \alpha_{n-k} W_s^2 ds$$
$$L_T^{(n+1)\perp} = c_n'' - \sum_{k=0}^{n} \int_0^T (T-s)^{2n+2} \beta_k \beta_{n-k} W_s^2 ds,\ n \geq 0.$$

Taking stochastic derivatives $D_t, D_t^{\perp}$ and conditional expectations on both sides we get

$$\begin{aligned}(T-s)^{2n+3} \alpha_n W_t &= 2 \sum_{k=0}^{n} \alpha_k \alpha_{n-k} W_t \int_t^T (T-s)^{2n+2} ds \\ &= \frac{2}{2n+3} W_t (T-t)^{2n+3} \sum_{k=0}^{n} \alpha_k \alpha_{n-k}, \\ (T-s)^{2n+3} \beta_n W_t^{\perp} &= -\frac{2}{2n+3} W_t^{\perp} (T-t)^{2n+3} \sum_{k=0}^{n} \beta_k \beta_{n-k},\end{aligned}$$

which means that

$$\alpha_{n+1} = \frac{2}{2n+3} \sum_{k=0}^{n} \alpha_k \alpha_{n-k},\ \beta_{n+1} = -\frac{2}{2n+3} \sum_{k=0}^{n} \beta_k \beta_{n-k}, n \geq 0.$$

Introducing $\alpha(s) = \sum_{n=0}^{\infty} \alpha_n s^{2n+1},\ \beta(s) = \sum_{n=0}^{\infty} \beta_n s^{2n+1}$ one obtains

$$\alpha'(s) = \alpha_0 + \sum_{n=0}^{\infty} (2n+3)\alpha_{n+1} s^{2n+2}$$

$$= 1 + 2\sum_{n=0}^{\infty}\sum_{k=0}^{n} (\alpha_k \alpha_{n-k}) s^{2n+2} = 1 + 2a^2(s),$$

$$\beta'(s) = \beta_0 + \sum_{n=0}^{\infty} (2n+3)\beta_{n+1} s^{2n+2}$$

$$= 1 - 2\sum_{n=0}^{\infty}\sum_{k=0}^{n} \beta_k \beta_{n-k} s^{2n+2} = 1 - 2\beta^2(s).$$

I.e.

$$\begin{aligned} \alpha'(s) &= 1 + 2a^2(s),\ \alpha(0) = 0, \\ \beta'(s) &= 1 - 2\beta^2(s),\ \beta(0) = 0. \end{aligned} \tag{11}$$

Thus

$$\alpha(s) = \frac{1}{\sqrt{2}} \tan(\sqrt{2}s),\ \beta(s) = -\frac{1}{\sqrt{2}} \tanh(\sqrt{2}s).$$

If $T < \frac{\pi}{2\sqrt{2}}$ series are convergent (not in BMO-space) and $(c, L, L^{\perp})$ is defined as $c = \frac{1}{2} \ln \cos(\sqrt{2}T)\cosh(\sqrt{2}T)$ (by calculations in the appendix),

$$L_t = \frac{1}{\sqrt{2}} \int_0^t \tan(\sqrt{2}s) W_s dW_s,\ L_t^{\perp} = -\frac{1}{\sqrt{2}} \int_0^t \tanh(\sqrt{2}s) W_s^{\perp} W_s^{\perp}.$$

When $T > \frac{\pi}{2\sqrt{2}}$ a local martingale L satisfying $L_T - \langle L \rangle_T = \frac{1}{2} \int_0^T W_t^2 dt$ does not exist (despite the fact that $\int_0^T W_t^2 dt$ is p-integrable for each $p \geq 1$), since from $\mathcal{E}_T(2L) = e^{\int_0^T W_t^2 dt}$ follows that $Ee^{\int_0^T W_t^2 dt} = E\mathcal{E}_T(2L) \leq 1$, which contradicts to $Ee^{\int_0^T W_t^2 dt} = \infty$ (see appendix).

In the next example exact solution of (5) also exists, however it does not belong to the extreme cases considered in [9],[10].

Example 2. Let $\alpha = -1,\ \gamma = 2,\ \bar{\eta} = 0,\ A_t = \int_0^t W_s W_s^{\perp} ds,\ \ \mathbf{F} = (\mathcal{F}_t^{W,W^{\perp}})$, where $W, W^{\perp}$ is a 2-dimensional Brownian motion. Then (5) becomes

$$L_T + L_T^{\perp} = c + \langle L \rangle_T - \langle L^{\perp} \rangle_T + \int_0^T W_s W_s^{\perp} ds.$$

We have

$$L_T^{(0)} = EL_T^{(0)} + \int_0^T (T-s)W_s^{\perp} dW_s, \ L_T^{(0),\perp} = EL_T^{(0),\perp} + \int_0^T (T-s)W_s dW_s^{\perp},$$

$$L_T^{(n+1)} + L_T^{(n+1)\perp} = c_n + \sum_{k=0}^{n} \langle L^{(k)}, L^{(n-k)} \rangle_T - \sum_{k=0}^{n} \langle L^{(k)\perp}, L^{(n-k)\perp} \rangle_T, \ n \geq 0.$$

We assert that

$$L_T^{(n)} = EL_T^{(n)} + \int_0^T (T-s)^{2n+1} (\alpha_n W_t + \beta_n W_s^{\perp}) dW_s,$$

$$L_T^{(n)\perp} = EL_T^{(n)\perp} + \int_0^T (T-s)^{2n+1} (\beta_n W_t - \alpha_n W_s^{\perp}) dW_s^{\perp},$$

where $\alpha_0 = 0, \ \ \beta_0 = 1$ and

$$\alpha_{n+1} = \frac{2}{2n+3} \sum_{k=0}^{n} (\alpha_k \alpha_{n-k} - \beta_k \beta_{n-k}), \ \beta_{n+1} = \frac{4}{2n+3} \sum_{k=0}^{n} \alpha_k \beta_{n-k}, n \geq 0.$$

Indeed,

$$\begin{aligned} L_T^{(n+1)} + L_T^{(n+1)\perp} &= c_n \\ + \sum_{k=0}^{n} \int_0^T (T-s)^{2n+2} (\alpha_k W_s + \beta_k W_s^{\perp})(\alpha_{n-k} W_s + \beta_{n-k} W_s^{\perp}) ds \\ - \sum_{k=0}^{n} \int_0^T (T-s)^{2n+2} (\beta_k W_s - \alpha_k W_s^{\perp})(\beta_{n-k} W_s - \alpha_{n-k} W_s^{\perp}) ds \\ = \sum_{k=0}^{n} \int_0^T (T-s)^{2n+2} [(\alpha_k \alpha_{n-k} - \beta_k \beta_{n-k}) W_s^2 - (\alpha_k \alpha_{n-k} - \beta_k \beta_{n-k}) W_s^{\perp 2} \\ + 2(\alpha_k \beta_{n-k} + \beta_k \alpha_{n-k}) W_s W_s^{\perp}] ds + c_n, \ n \geq 0. \end{aligned}$$

Using representation of integrands by stochastic derivatives we get

$$
\begin{aligned}
&(T-t)^{2n+3}(\alpha_{n+1}W_t+\beta_{n+1}W_t^{\perp})\\
&=E[D_t(\sum_{k=0}^{n}\langle L^{(k)},L^{(n-k)}\rangle_T-\sum_{k=0}^{n}\langle L^{(k)\perp},L^{(n-k)\perp}\rangle_T)|\mathcal{F}_t]\\
&=2\sum_{k=0}^{n}[(\alpha_k\alpha_{n-k}-\beta_k\beta_{n-k})W_t+(\alpha_k\beta_{n-k}+\beta_k\alpha_{n-k})W_t^{\perp}]\int_t^T(T-s)^{2n+2}ds\\
&=\frac{2(T-t)^{2n+3}}{2n+3}\sum_{k=0}^{n}[(\alpha_k\alpha_{n-k}-\beta_k\beta_{n-k})W_t+(\alpha_k\beta_{n-k}+\beta_k\alpha_{n-k})W_t^{\perp}],\\
&(T-t)^{2n+3}(\beta_{n+1}W_t-\alpha_{n+1}W_t^{\perp})\\
&=E[D_t^{\perp}(\sum_{k=0}^{n}\langle L^{(k)},L^{(n-k)}\rangle_T-\sum_{k=0}^{n}\langle L^{(k)\perp},L^{(n-k)\perp}\rangle_T)|\mathcal{F}_t]\\
&=2\sum_{k=0}^{n}[-(\alpha_k\alpha_{n-k}-\beta_k\beta_{n-k})W_t^{\perp}+(\alpha_k\beta_{n-k}+\beta_k\alpha_{n-k})W_t]\int_t^T(T-s)^{2n+2}ds\\
&=\frac{2(T-t)^{2n+3}}{2n+3}\sum_{k=0}^{n}[-(\alpha_k\alpha_{n-k}-\beta_k\beta_{n-k})W_t^{\perp}+(\alpha_k\beta_{n-k}+\beta_k\alpha_{n-k})W_t].
\end{aligned}
$$

Equalising coefficients at $W, W^{\perp}$ we obtain the desired formula. One can be checked that $\lim_{n\to\infty}\sqrt[n]{|a_n|}=0, \lim_{n\to\infty}\sqrt[n]{|b_n|}=0$. Introducing $\alpha(s)=\sum_{n=0}^{\infty}\alpha_n s^{2n+1},\ \beta(s)=\sum_{n=0}^{\infty}\beta_n s^{2n+1}$ one obtains

$$
\begin{aligned}
L_t &= L_0+\int_0^t(\alpha(T-s)W_s+\beta(T-s)W_s^{\perp})dW_s,\\
L_t^{\perp} &= L_0^{\perp}+\int_0^t(\beta(T-s)W_s-\alpha(T-s)W_s^{\perp})dW_s^{\perp}.
\end{aligned}
$$

On the other hand we can derive ODE for the pair (α,β)

$$
\begin{aligned}
\alpha'(s) &= 2\alpha^2(s)-2\beta^2(s),\ \alpha(0)=0,\\
\beta'(s) &= 1+4\alpha(s)\beta(s),\ \beta(0)=0.
\end{aligned}
\tag{12}
$$

Indeed

$$\alpha'(s) = \alpha_0 + \sum_{n=0}^{\infty}(2n+3)\alpha_{n+1}s^{2n+2}$$

$$= 2\sum_{n=0}^{\infty}\sum_{k=0}^{n}(\alpha_k\alpha_{n-k} - \beta_k\beta_{n-k})s^{2n+2} = 2a^2(s) - 2\beta^2(s),$$

$$\beta'(s) = \beta_0 + \sum_{n=0}^{\infty}(2n+3)\beta_{n+1}s^{2n+2}$$

$$= 1 + 4\sum_{n=0}^{\infty}\sum_{k=0}^{n}\alpha_k\beta_{n-k}s^{2n+2} = 1 + 4\alpha(s)\beta(s).$$

The equation (12) is easy to solve, if we pass to the equation for complex-variable function $\zeta(s) = \alpha(s) + i\beta(s)$

$$\zeta'(s) = i + 2\zeta^2(s),\ \zeta(0) = 0.$$

It is obvious that $\zeta(s) = \frac{1}{1-i}\tan((1+i)s)$ is a solution. We have

$$\zeta(s) = \frac{1}{2}(1+i)\frac{\sin((1+i)s)\cos((1-i)s)}{|\cos((1+i)s)|^2}$$

$$= \frac{1}{4}(1+i)\frac{\sin(2s) + i\sinh(2s)}{|\cos((1+i)s)|^2}$$

$$= \frac{1}{4}\frac{\sin(2s) - \sinh(2s) + i(\sin(2s) + \sinh(2s))}{\cos^2(s)\cosh^2(s) + \sin^2(s)\sinh^2(s)}.$$

Finally we can write explicit solution

$$\alpha(s) = \frac{1}{4}\frac{\sin(2s) - \sinh(2s)}{\cos^2(s)\cosh^2(s) + \sin^2(s)\sinh^2(s)},$$

$$\beta(s) = \frac{1}{4}\frac{\sin(2s) + \sinh(2s)}{\cos^2(s)\cosh^2(s) + \sin^2(s)\sinh^2(s)}$$

of (12) and conclude that it exists on whole $[0,\infty)$, since the denominator does not vanish.

A Appendix

The formula $Ee^{-T^2\int_0^1 W_t^2 dt} = \frac{1}{\sqrt{\cosh(\sqrt{2}T)}}$ is derived in [7]. Similarly we can prove

Proposition 3.

$$Ee^{\int_0^T W_t^2 dt} = \begin{cases} \frac{1}{\sqrt{\cos(\sqrt{2}T)}}, \text{ if } T < \frac{\pi}{2\sqrt{2}} \\ \infty, \text{ if } T \geq \frac{\pi}{2\sqrt{2}} \end{cases} .$$

Proof. Let $e_n(t)$ be orhonormal basis in $L^2[0,1]$. Then $Ee^{\int_0^T W_t^2 dt} = Ee^{T^2 \int_0^1 W_t^2 dt} = Ee^{T^2 \sum_{n=1}^{\infty} (\int_0^1 e_n(t) W_t dt)^2} = E \prod_{n=1}^{\infty} e^{T^2 (\int_0^1 e_n(t) W_t dt)^2}$. Since

$$E(\int_0^1 e_n(t) W_t dt)(\int_0^1 e_m(t) W_t dt) = \int_0^T e_n(t) \int_0^T (t \wedge s) e_m(s) ds dt$$

it is convenient to use the orthonormal basis of eigenvectors of the operator $\int_0^T (t \wedge s) f(s) ds$ in $L^2[0,1]$. From $\lambda f(t) = \int_0^T (t \wedge s) f(s) ds$ follows that $\lambda f''(t) = -f(t),\ f(0) = 0,\ f'(1) = 0$. The function $\sin \mu \pi t$ satisfies these conditions iff $\mu^2 = 1/\lambda$, $\cos \mu\pi = 0$ and $\mu = -1/2 + n$. Thus

$$\lambda_n = \frac{1}{(n-1/2)^2 \pi^2},\ e_n(t) = \sqrt{2} \sin((n-1/2)\pi t),\ n \geq 1$$

and $E(\int_0^1 e_n(t) W_t dt)(\int_0^1 e_m(t) W_t dt) = \lambda_n \int_0^1 e_n(t) e_m(t) dt = 0,\ n \neq m$. Since random variables $(\int_0^1 e_n(t) W_t dt)$ are orthogonal and normal they are also independent. Hence taking into account infinite product decomposition of $\cos(\sqrt{2}t)$ one gets

$$Ee^{\int_0^T W_t^2 dt} = \prod_{n=1}^{\infty} Ee^{T^2 (\int_0^1 e_n(t) W_t dt)^2}$$

$$= \prod_{n=1}^{\infty} Ee^{T^2 \lambda_n W_1^2} = \prod_{n=1}^{\infty} \frac{1}{\sqrt{1 - \frac{2T^2}{(n-1/2)^2 \pi^2}}}$$

$$= \sqrt{\prod_{n=1}^{\infty} \frac{1}{1 - \frac{8T^2}{(2n-1)^2 \pi^2}}} = \frac{1}{\sqrt{\cos(\sqrt{2}T)}},$$

if $\sqrt{2}T < \pi/2$.
It easy to see that

$$E \exp\left(\int_0^{\frac{\pi}{2\sqrt{2}}} W_t^2 dt\right) = \lim_{T \uparrow \frac{\pi}{2\sqrt{2}}} E \exp\left(\int_0^T W_t^2 dt\right) = \lim_{T \uparrow \frac{\pi}{2\sqrt{2}}} \frac{1}{\sqrt{\cos(\sqrt{2}T)}} = \infty.$$

If $T > \frac{\pi}{2\sqrt{2}}$ then $Ee^{\int_0^T W_t^2 dt} > Ee^{\int_0^{\frac{\pi}{2\sqrt{2}}} W_t^2 dt} = \infty$.

Lemma 3. *Let $(a_n)_{n\geq 0}$ be a solution of the system*

$$a_0 = 1,\ a_{n+1} = \sum_{k=0}^{n} a_k a_{n-k}. \tag{13}$$

Then $a_n = \frac{1}{4n+2}\binom{2n+2}{n+1}$.

Proof. For the series $u(\lambda) = \sum_{n=0}^{\infty} a_n \lambda^n$ from (13) we get equation $u(\lambda) = 1 + \lambda u^2(\lambda)$, with the roots $u(\lambda) = \frac{1}{2\lambda}(1 \pm \sqrt{1-4\lambda})$. The equality $u(\lambda) = \frac{1}{2\lambda}(1 + \sqrt{1-4\lambda})$ is impossible, since decomposition of the right hand side is starting from the term $\frac{1}{\lambda}$. Therefore, equality $a_n = \frac{1}{4n+2}\binom{2n+2}{n+1}$ follows from the Taylor expansion of $1 - \sqrt{1-4\lambda}$, since

$$\begin{aligned} u(\lambda) &= \frac{1}{2\lambda}(1 - \sqrt{1-4\lambda}) \\ &= -\frac{1}{2}\sum_{n\geq 1} \frac{\frac{1}{2}(\frac{1}{2}-1)\cdots(\frac{1}{2}-n+1)}{n!}(-4)^n \lambda^{n-1} \\ &= \frac{1}{2}\sum_{n\geq 1} \frac{(2-1)\cdots(2n-2-1)}{2^n n!} 4^n \lambda^{n-1} \\ = \frac{1}{2}\sum_{n\geq 1} \frac{(2n-3)!!}{n!} 2^n \lambda^{n-1} &= \frac{1}{2}\sum_{n\geq 1} \frac{1}{2n-1}\binom{2n}{n}\lambda^{n-1}. \end{aligned}$$

Lemma 4. *There exist sequences $(m_i, i \geq 1) \in \mathcal{M}$, $(m_i^\perp, i \geq 1) \in \mathcal{M}^\perp$, such that $e^\eta = c_1 \frac{\mathcal{E}_T(m_1)}{\mathcal{E}_T(m_1^\perp)} \mathcal{E}_T^2(m_1^\perp)$ and*

$$e^\eta = c_n \frac{\mathcal{E}_T(\sum_i^n m_i)}{\mathcal{E}_T(\sum_i^n m_i^\perp)} \mathcal{E}_T^2(m_n'^\perp),\ n \geq 2, \tag{14}$$

where $m_n'^\perp = m_n^\perp - \langle m_n^\perp, \sum_i^{n-1} m_i^\perp \rangle$.

Proof. The theorem will be proved by induction. Assume (14) is valid for n. There exist such martingales $m_{n+1}, m_{n+1}^\perp$ that $c'\mathcal{E}_T(m_{n+1}' + m_{n+1}'^\perp) = \mathcal{E}_T^2(m_n'^\perp)$ and

$$m_{n+1}' = m_{n+1} - \langle m_{n+1}, \sum_i^n, m_i \rangle,\quad m_{n+1}'^\perp = m_{n+1}^\perp - \langle m_{n+1}^\perp, \sum_i^n m_i^\perp \rangle$$

are martingales w.r.t. $\mathcal{E}(\sum_i^n m_i + m_i^{\perp}) \cdot P$. Thus

$$e^{\eta} = c_n c' \frac{\mathcal{E}_T(\sum_i^n m_i)}{\mathcal{E}_T(\sum_i^n m_i^{\perp})} \mathcal{E}(m_{n+1} - \langle m_{n+1}, \sum_i^n m_i \rangle + m_{n+1}^{\perp} - \langle m_{n+1}^{\perp}, \sum_i^n m_i^{\perp} \rangle)$$

$$= c_{n+1} \frac{\mathcal{E}_T(\sum_i^n m_i)\mathcal{E}_T(m_{n+1} - \langle m_{n+1}, \sum_i^n m_i \rangle)}{\mathcal{E}_T(\sum_i^n m_i^{\perp})\mathcal{E}_T(m_{n+1}^{\perp} - \langle m_{n+1}^{\perp}, \sum_i^n m_i^{\perp} \rangle)} \mathcal{E}_T^2(m_{n+1}^{\perp} - \langle m_{n+1}^{\perp}, \sum_i^n m_i^{\perp} \rangle)$$

$$= c_{n+1} \frac{\mathcal{E}_T(\sum_i^{n+1} m_i)}{\mathcal{E}_T(\sum_i^{n+1} m_i^{\perp})} \mathcal{E}_T^2(m_{n+1}'^{\perp}).$$

Remark. If we will prove the convergence of series $\sum_i m_i, \sum_i m_i^{\perp}$, then $m_n^{\perp} \to 0, m_n'^{\perp} \to 0,\ \mathcal{E}(m_n'^{\perp}) \to 1$ and $e^{\eta} = c\frac{\mathcal{E}_T(\sum_i^{\infty} m_i)}{\mathcal{E}_T(\sum_i^{\infty} m_i^{\perp})}$.

References

[1] G. Adomian, Nonlinear Stochastic Systems Theory and Applications to Physics, Kluwer, (1989).

[2] F. Biagini, P. Guasoni and M. Pratelli, Mean variance hedging for stochastic volatil- ity models,Math. Finance, Vol. 10. no. 2, 2000,109-123

[3] L. Gabet, The Theoretical Foundation of the Adomian Method, Computers Math. Applic. Vol. 27, No. 12, (1994), 41-52,

[4] M. De Donno, P. Guasoni, M. Pratelli, Superreplication and Utility Maximization in Large Financial Markets, Stochastic Processes and their Applications, 115 (2005), no. 12, 2006-2022,

[5] C. Dellacherie and P. A. Meyer, Probabilit´es et potentiel. Chapitres V a VIII. Th´eorie des martingales. Actualit´es Scientifiques et Industrielles Hermann, Paris, (1980).

[6] J. Jacod, Calcule Stochastique et probl‘emes des martingales. Lecture Notes in Math., Springer, Berlin, N. 714, (1979).

[7] H-H. Kuo, Gaussian measures in Banach spacese, Lecture notes in Mathematics, Springer, Berlin, 463, (1975).

[8] M. Mania and R. Tevzadze, An exponential martingale equation , Electronic Communications in Probability 11, (2006), 206–216.

[9] M. Mania and R. Tevzadze, A Semimartingale Bellman equation and the variance-optimal martingale measure, Georgian Math. J. vol. 7, No 4 (2000), 765–792.

[10] M. Mania and R. Tevzadze, A Semimartingale Backward Equation and the Variance optimal martingale measure under general information flow, SIAM Journal on Control and Optimization, Vol. 42, N5, (2003), 1703-1726.

[11] R. Tevzadze, Solvability of Backward Stochastic Differential Equation with Quadratic Growth, Stochastic Processes and their Applications, vol. 118, №3, (2008), 503-515.

Number of Unordered Samples of Integers With a Given Sum

Tsotne Kutalia

Cybernetics Institute of Georgian Technical University.

Abstract

There is an analytic formula counting the number of ordered samples of $\boldsymbol{N}$ non-negative integers making up a given sum. In this paper we study the number of unordered samples of $\boldsymbol{N}$ non-negative integers with a given sum. We produce a closed form solution for $\boldsymbol{N} = \mathbf{3}$ non-negative integers.

Keywords: Combinatorics, Number Theory, Graph Theory

1 Introduction

A typical approach to finding the total number of ordered samples $(a_1+...+a_N)$ of N non-negative integers making up a sum of $n(n \geq N, a_1 + ...a_N = n)$ is to take n ones $1 + 1 + ... + 1$ (n times) and put $N - 1$ separator bars in the sequence. The total umber of arrangements of bars and ones can be viewed as the total number of ordered arrangements of $N - 1$ zeros and n ones which obviously is C^n_{n+N-1}. See [1].

However, the same problem gets complicated for unordered samples. There is a known recursion in ([2]) which is defined as

$$f_N(n) = f_{N-1}(n) + f_N(n - N) \tag{1}$$

In the text that follows, we obtain a precise formula for $N = 3$ and $n \geq N$ to be

$$f_3(n) = I_{\{\{n\}_3=0\}}\left[\frac{(n+3)(n+6)}{18} + I_{\{\{n\}_2=0\}}\frac{n^2}{36} + I_{\{\{n\}_2\neq 0\}}\frac{(n-3)(n+3)}{36}\right] +$$
$$I_{\{\{n\}_3=1\}}\left[\frac{(n+2)(n+5)}{18} + I_{\{\{n\}_2=0\}}\frac{(n+2)^2}{36} + I_{\{\{n\}_2\neq 0\}}\frac{(n-1)(n+5)}{36}\right] +$$
$$I_{\{\{n\}_3=2\}}\left[\frac{(n+1)(n+4)}{18} + I_{\{\{n\}_2=0)\}}\frac{(n+4)^2}{36} + I_{\{\{n\}_2\neq 0\}}\frac{(n+1)(n+7)}{36}\right] \tag{2}$$

which reduces to

$$f_3(n) = \frac{(n+3-\{n\}_3)(n+6-\{n\}_3)}{18} + \frac{(n+2(\{n\}_3)^2) - (3\{n\}_3)^2}{36} \tag{3}$$

where $\{n\}_k$ denotes a modulo operator giving a remainder for division of n over k.

2 Graphical Representation of Partitions, N=3

Let us denote by $f_N(n)$ the function counting the number of unordered samples of N non-negative integers $[a_1, ..., a_N]$ such that $a_1 + ... + a_N = n$. Let by convention $f_0(n) = 1$ for all n. Obviously $f_1(n) = 1$ for all n as well. It can easily be checked that for even $n, f_2(n) = \frac{n+2}{2}$ and for odd n we have $f_2(n) = \frac{n+1}{2}$. We can thus define $f_2(n)$ with the indicator functions as

$$f_2(n) = I_{\{n \ mod \ 2=0\}}\frac{n+2}{2} + I_{\{n \ mod \ 2=1\}}\frac{n+1}{2} \tag{4}$$

For $N = 3$, we take the sum of N ones and partition the sum of series with 2 separator bars. This can best be illustrated through an example. For $n = 3$ we have the following arrangements of 2 separator bars

$$|| \, 1+1+1, \qquad | \, 1 \, | +1+1, \qquad 1 \, | +1 \, | +1 \tag{5}$$

The first arrangement in (5) corresponds to $a_1 = 0, a_2 = 0, a_3 = 3$. The second arrangement corresponds to $a_1 = 0, a_2 = 1, a_3 = 2$ and the last one to $a_1 = 1, a_2 = 1, a_3 = 1$. So we have the followig sample $\{003, 012, 111\}$. Note that the numbers in each sample are listed in a non-decreasing order. That is why the arrangement like $| \, 1+1 \, | +1$ are ignored since that would correspond to the sample element 021 in which the numbers are not put in non-decreasing order and thus such an element already extists as 012.
We can enumerate the positions of separator bars in the series of ones as follows $^1 1^2 + 1^3 + 1^4$ where the superscripts mark the positions of possible placements of the separator bars. Then the sample $\{003, 012, 111\}$ can be transformed into the following sample $\{11, 12, 23\}$. In this sample, the first element 11

stands for the two bars placed at the position 1 and thus it corresponds to the first partition in (5). The element 12 corresponds to the second partition and respectively 23 is for the third partition.
We view the sample elements as the coordinates of points on the cartesian coordinate system and for convenience we reverse the numbers. So $\{11, 12, 23\}$ becomes $\{11, 21, 32\}$. The points on the coordinate system corresponding to this sample is

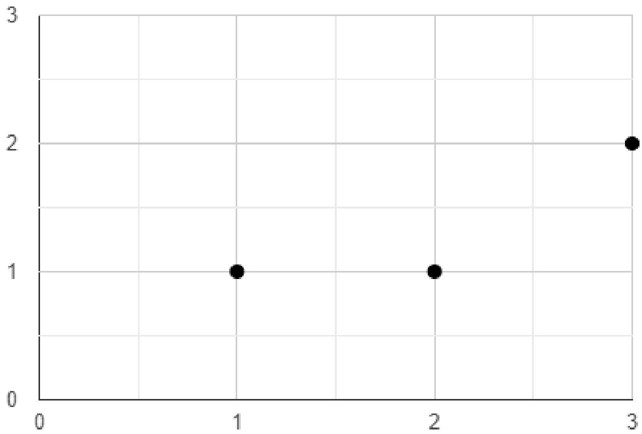


Fig. 1: $f_3(3) = 3$

Likewise, for $n = 4$ we have the following partitions identical to (5) (the corresponding samples and the reverse versions of them are given below each partition)

$$|| \, 1+1+1+1, \quad | \, 1 \, | +1+1+1, \quad 1 \, | +1 \, | +1+1, \quad 1 \, | +1 \, | +1+1 \tag{6}$$

Partition:	$\|\| \, 1+1+1+1$	$\| \, 1 \, \| +1+1+1$	$1 \, \| +1 \, \| +1+1$	$1 \, \| +1 \, \| +1+1$
Sample:	004	013	022	112
Coordinate:	11	12	13	23
Reversed:	11	21	31	32

and the corresponding plot for $f_3(4) = 4$ is on Fig.2 below.

The appendix at the end of the paper contains some of the partitions and the respective graphs. We take some of the examples here to develop the

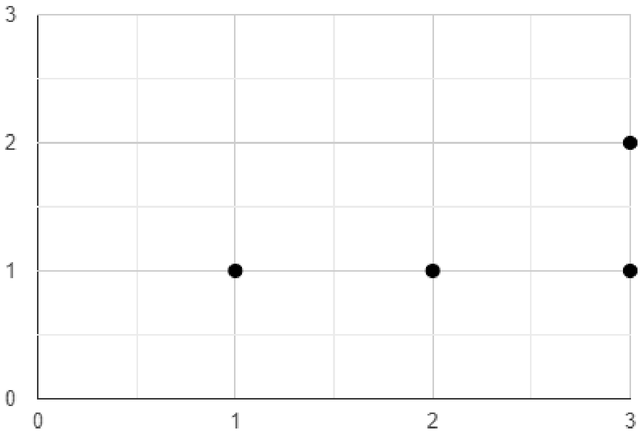


Fig. 2: $f_3(4) = 4$

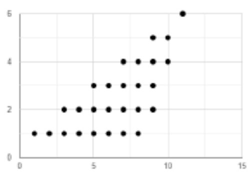


Fig. 3: $f_3(15) = 27$

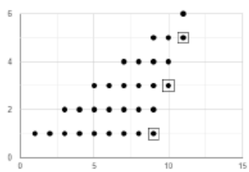


Fig. 4: $f_3(16) = 30$

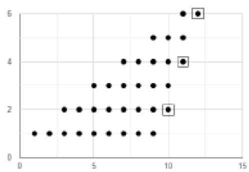


Fig. 5: $f_3(17) = 33$

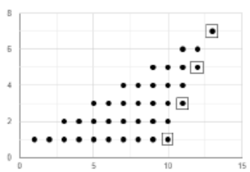


Fig. 6: $f_3(18) = 37$

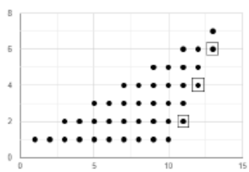


Fig. 7: $f_3(19) = 40$

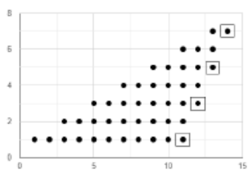


Fig. 8: $f_3(20) = 44$

formula (2). The examples for $N = 3$ are $n = 15, n = 16, n = 17, n = 18, n = 19, n = 20$.
Let us begin with $n = 15, n = 16$ and $n = 17$ on the one hand and $n = 18, n =$

$19, n = 20$ on another. The respective graphs are given in Fig. 3 to Fig. 8. There are some interesting patters emerging. In particular, we have 3 possible configurations listed below

Configuration 1, $n \ mod \ 3 = 0$: Fig.3 displays the case when $n = 15$ which is divisible by 3. On that graph there is an extreme point placed at the coordinate $(11, 6)$. This point is unique in the sense that it does not share either x or y coordinate with any other point. In general, there is a point located at the coordinate (x_0, y_0) while there is no any other point having either x_0 as x coordinate or y_0 as y coordinate. The value of y_0 coordinate can be found by

$$y_0 = 1 + \frac{n}{3} \tag{7}$$

Configuration 2, $n \ mod \ 3 = 1$: Fig. 4 displays the case when $n = 16$. The points put in squares indicate the additions to the previous graph. So as we move from Fig. 3 to Fig. 4 we have new points added on the coordinates $(11, 5)$, $(10, 3)$ and $(9, 1)$. In general, we have the points added on the coordinates $(x_0, y_0 - 1)$, $(x_0 - 1, y_0 - 3)$ and so on till the last y coordinate reaches 1. i.e. $y = 1$. The value of y_0 coordinate now is

$$y_0 = 1 + \frac{n-1}{3} \tag{8}$$

Configuration 3, $n \ mod \ 3 = 2$: Fig. 5 displays the case when $n = 17$. Again, the points in the squares indicate the additions from the previous case. In particular when moving from Fig. 4 to Fig. 5 we have the new points added on the coordinates $(12, 6)$, $(11, 4)$ and $(10, 2)$. In general, the points are added on $(x_0 + 1, y_0)$, $(x_0, y_0 - 2)$ and so on till the last point's y coodinate reaches 2. The value of y_0 for this configuration is

$$y_0 = 1 + \frac{n-2}{3} \tag{9}$$

In total we only have these 3 configurations and the cycle goes over and over again. For example, when $n = 18$, the configuration is similar to the case when $n = 15$. In general, all $n \ mod \ 3 = 0$ configurations are similar with a slight difference. When n is odd, the last added point occurs at the coordinate $y = 2$ while for even n, the additions continue till $y = 1$. Similar differences hold for cases $n \ mod \ 3 = 1$ and $n \ mod \ 3 = 2$. In particular, for odd $n - 1$, we keep adding points as described in Configuration 2 till the last point's y coordinate is $y = 1$ while for even $n - 1$, the last point added occurs at the y coordinate of $y = 2$. Similarly for odd $n - 2$ we have the last added point's y coordinate to be $y = 2$ and for even $n - 2$ we have the same coordinate to be $y = 1$.

$f_3(n)$ is simply the number of points on a corresponding plot. In order to count them we take the diagonal approach. Let us observe the counting method for all 3 configurations for the above mentioned examples.

Configuration 1: We can split the total number of points on Fig. 3 into two parts. The upper part of (and including) the main diagonal and the lower part.

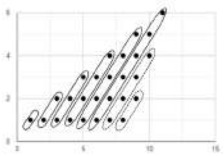


Fig. 9: $f_3(15) = 27$

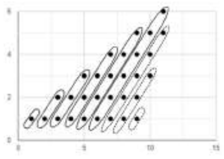


Fig. 10: $f_3(16) = 30$

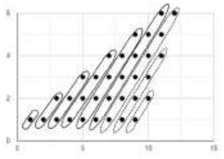


Fig. 11: $f_3(17) = 33$

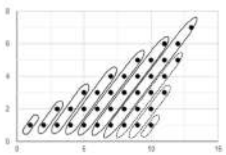


Fig. 12: $f_3(18) = 37$

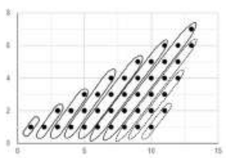


Fig. 13: $f_3(19) = 40$

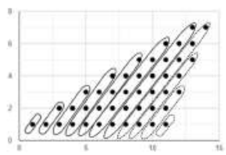


Fig. 14: $f_3(20) = 44$

We refer to the formula of the sum n terms of arithmetic series which in its more convenient form can according to [REFERENCE HERE] be written as

$$S_n = \frac{n}{2}(a_1 + a_n) \tag{10}$$

where a_1 and a_n are respectively the first and the last terms of the series.
In configuration 1, the upper part of (and including) the longest diagonal is summed as $1 + 2 + ... + (1 + \frac{n}{3})$ where the last term comes from (7). By (10) this sum is $\frac{(n+3)(n+6)}{18}$. As for the lower part of the diagonal, we have 2 variations. In particular, when n is odd (the case shown on Fig. 9) the sum of the arithmetic series with the common difference of 2 consisting of the following terms $1 + 3 + 5 + ... + (\frac{n}{3} - 1)$ which by (10) is $\frac{n^2}{36}$. On the other hand, if n is even (the case shown on Fig. 12), the sum of the arithmetic series is $2 + 4 + 6 + ... + (\frac{n}{3} - 1)$ which by (10) is $\frac{(n-3)(n+3)}{36}$. Combining these terms yields the number of unordered samples of 3 non - negative integers with a sum n when $n \ mod \ 3 = 0$ which is

$$I_{\{\{n\}_3=0\}} \left[\frac{(n+3)(n+6)}{18} + I_{\{\{n\}_2=0\}} \frac{n^2}{36} + I_{\{\{n\}_2\neq 0\}} \frac{(n-3)(n+3)}{36} \right]. \tag{11}$$

Similarly, for configuration 2, the upper part of (and including) the longest diagonal is summed as $1 + 2 + ... + (1 + \frac{n-1}{3})$ which by (10) is $\frac{(n+2)(n+5)}{18}$. The lower parts differ according to whether $n - 1$ is odd or even. For odd $n - 1$ (the case shown on Fig. 10) the sum of the arithmetic series is $2 + 4 + 6 +$

$... + \frac{n-1}{3}$ which by (10) is $\frac{(n-1)(n+5)}{18}$. For $n-1$ being even, the series becomes $1+3+5+...+\frac{n-1}{3}$ which by (10) is $\frac{(n+2)^2}{36}$. Combininig these terms yields the number of unordered samples of 3 non - negative integers with a sum n when $n \ mod \ 3 = 1$ which is

$$I_{\{\{n\}_3=1\}}\left[\frac{(n+2)(n+5)}{18} + I_{\{\{n\}_2=0\}}\frac{(n+2)^2}{36} + I_{\{\{n\}_2\neq 0\}}\frac{(n-1)(n+5)}{36}\right]. \tag{12}$$

Finally, for configuration 3, the upper part of (and including) the lonest diagonal is summed as $1+2+3+...+(1+\frac{n-3}{3})$ which by (10) is $\frac{(n+1)(n+4)}{38}$. The lower parts similarly to the previous configurations is differ according to $n-2$ being odd or even. For odd $n-2$, the sum is $2+4+...+\frac{n-1}{3}$ which by (10) is $\frac{(n+4)^2}{36}$ and for odd $n-2$ the sum $1+3+...+(1+\frac{n-3}{3})$ by (10) is $\frac{(n+1)(n+7)}{36}$. Combininig these terms yields the number of unordered samples of 3 non - negative integers with a sum n when $n \ mod \ 3 = 2$ which is

$$I_{\{\{n\}_3=2\}}\left[\frac{(n+1)(n+4)}{18} + I_{\{\{n\}_2=0\}}\frac{(n+4)^2}{36} + I_{\{\{n\}_2\neq 0\}}\frac{(n+1)(n+7)}{36}\right] \tag{13}$$

In total, $f_3(n)$ turns out to be the sum of (11), (12) and (13) which is (2) restated below

$$\begin{aligned} f_3(n) = I_{\{\{n\}_3=0\}}&\left[\frac{(n+3)(n+6)}{18} + I_{\{\{n\}_2=0\}}\frac{n^2}{36} + I_{\{\{n\}_2\neq 0\}}\frac{(n-3)(n+3)}{36}\right] + \\ I_{\{\{n\}_3=1\}}&\left[\frac{(n+2)(n+5)}{18} + I_{\{\{n\}_2=0\}}\frac{(n+2)^2}{36} + I_{\{\{n\}_2\neq 0\}}\frac{(n-1)(n+5)}{36}\right] + \\ I_{\{\{n\}_3=2\}}&\left[\frac{(n+1)(n+4)}{18} + I_{\{\{n\}_2=0)\}}\frac{(n+4)^2}{36} + I_{\{\{n\}_2\neq 0\}}\frac{(n+1)(n+7)}{36}\right] \end{aligned}$$

It is easily verified that the formula above can be reduced to (3) component by component. This is also restated below

$$f_3(n) = \frac{(n+3-\{n\}_3)(n+6-\{n\}_3)}{18} + \frac{(n+2(\{n\}_3)^2)-(3\{n\}_3)^2}{36}$$

At this point it remains to prove the formula. This is done by induction (1) part by part.

To prove that the configuration 1 part of the formula holds for any $n \geq 3$, we assume that it holds for some $n \geq 3$ and show that it also holds for $n+3$. In fact, it can easily be shown that if we put $n+3$ in place of n in (7), we obtain the following sum in place of (11)

$$\begin{aligned} I_{\{\{n+3\}_3=0\}}(1+2+3+...+(1+\frac{n+3}{3})+\\ I_{\{\{n+3\}_2=0\}}(1+3+5+...+(\frac{n+3}{3}-1))+\\ I_{\{\{n+3\}_2\neq 0\}}(2+4+6+...+(\frac{n+3}{3}-1))). \end{aligned} \tag{14}$$

Applying (10) to each component yields

$$I_{\{\{n+3\}_3=0\}}\left[\frac{6+n}{6}\frac{9+n}{3}+I_{\{\{n+3\}_2=0\}}\frac{(n+3)^2}{36}+I_{\{\{n+3\}_2\neq 0\}}\frac{n(n+6)}{36}\right]. \tag{15}$$

Similarly, for configuration 2 part of the formula, we get

$$\begin{aligned} I_{\{\{n+3\}_3=1\}}(1+2+3+...+(1+\frac{n+2}{3})+\\ I_{\{\{n+3\}_2=0\}}(2+4+6+...+\frac{n+2}{3})+\\ I_{\{\{n+3\}_2\neq 0\}}(1+3+5+...+\frac{n+2}{3})). \end{aligned} \tag{16}$$

This by applying (10) becomes

$$I_{\{\{n+3\}_3=1\}}\left[\frac{5+n}{6}\frac{9+n}{3}+I_{\{\{n+3\}_2=0\}}\frac{n+2}{12}\frac{8+n}{3}+I_{\{\{n+3\}_2\neq 0\}}\frac{(n+5)^2)}{36}\right]. \tag{17}$$

Lastly, for configuration 3 part of the formula, we have

$$\begin{aligned} I_{\{\{n+3\}_3=2\}}(1+2+3+...+(1+\frac{n+1}{3})+\\ I_{\{\{n+3\}_2=0\}}(1+3+5+...+(1+\frac{n+1}{3}))+\\ I_{\{\{n+3\}_2\neq 0\}}(2+4+6+...+(1+\frac{n+1}{3}))). \end{aligned} \tag{18}$$

This by applying (10) becomes

$$I_{\{\{n+3\}_3=1\}}\left[\frac{4+n}{6}\frac{7+n}{3}+I_{\{\{n+3\}_2=0\}}\frac{(n+7)^2)}{36}+I_{\{\{n+3\}_2\neq 0\}}\frac{4+n}{12}\frac{10+n}{3}\right]. \tag{19}$$

Combining (15), (17) and (19) yields $f_3(n+3)$ defined by (2).

On the other hand, by taking arbitrary non - negative integers, the correctness of (2) and (3) can be easily verified by (1).

3 General Recursive Formula for Arbitrary N and $n >= N$

In terms of modulus operators, (1) can be redefined for different N-s. For $N = 4$, we have

$$\begin{aligned} f_4(n) = I_{\{\{n\}_4=0\}} \sum_{k=1}^{\frac{n}{4}+1} f_{4k-4}(3) + I_{\{\{n\}_4=1\}} \sum_{k=1}^{\frac{n-1}{4}+1} f_{4k-3}(3) + \\ I_{\{\{n\}_4=2\}} \sum_{k=1}^{\frac{n-2}{4}+1} f_{4k-2}(3) + I_{\{\{n\}_4=3\}} \sum_{k=1}^{\frac{n-3}{4}+1} f_{4k-1}(3) \end{aligned} \tag{20}$$

For $N = 5$, we have

$$\begin{aligned} f_5(n) = I_{\{\{n\}_5=0\}} \sum_{k=1}^{\frac{n}{5}+1} f_{5k-5}(4) + I_{\{\{n\}_5=1\}} \sum_{k=1}^{\frac{n-1}{5}+1} f_{5k-4}(4) + \\ I_{\{\{n\}_5=2\}} \sum_{k=1}^{\frac{n-2}{5}+1} f_{5k-3}(4) + I_{\{\{n\}_5=3\}} \sum_{k=1}^{\frac{n-3}{5}+1} f_{5k-2}(4) + \\ I_{\{\{n\}_5=4\}} \sum_{k=1}^{\frac{n-4}{5}+1} f_{5k-1}(4) \end{aligned} \tag{21}$$

For $N = 6$, we have

$$\begin{aligned} f_6(n) = I_{\{\{n\}_6=0\}} \sum_{k=1}^{\frac{n}{6}+1} f_{6k-6}(5) + I_{\{\{n\}_6=1\}} \sum_{k=1}^{\frac{n-1}{6}+1} f_{6k-5}(5) + \\ I_{\{\{n\}_6=2\}} \sum_{k=1}^{\frac{n-2}{6}+1} f_{6k-4}(5) + I_{\{\{n\}_6=3\}} \sum_{k=1}^{\frac{n-3}{6}+1} f_{6k-3}(5) + \\ I_{\{\{n\}_6=4\}} \sum_{k=1}^{\frac{n-4}{6}+1} f_{6k-2}(5) + I_{\{\{n\}_6=5\}} \sum_{k=1}^{\frac{n-5}{6}+1} f_{6k-1}(5) \end{aligned} \tag{22}$$

In general, for an arbitrary N, we have (1)

$$f_N(n) = I_{\{\{n\}_N=0\}} \sum_{k=1}^{\frac{n}{N}+1} f_{Nk-N}(N-1) + I_{\{\{n\}_N=1\}} \sum_{k=1}^{\frac{n-1}{N}+1} f_{Nk-N+1}(N-1) + ...+ I_{\{\{n\}_N=N-1\}} \sum_{k=1}^{\frac{n-N+1}{N}+1} f_{Nk-1}(N-1) = \sum_{j=1}^{N-1} I_{\{\{n\}_N=j\}} \sum_{k=1}^{\frac{n-j}{N}+1} f_{Nk-N+j}(N-1) \tag{23}$$

References

[1] Shiryaev A.N., Problems in Probability, Springer, 2012, pp.4

[2] Shiryaev A.N., Erlikh I.G., Yaskov P.A., Probability in Theorems and Problems, pp. 12

Appendix A Scatter Configurations for $N = 3$

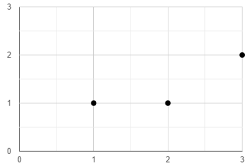


Fig. A1: $f_3(3) = 3$

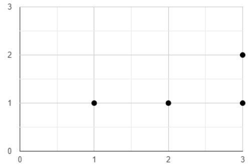


Fig. A2: $f_3(4) = 4$

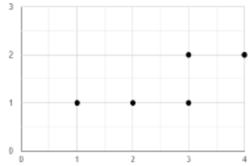


Fig. A3: $f_3(5) = 5$

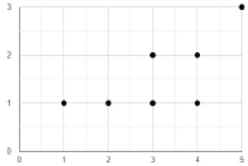


Fig. A4: $f_3(6) = 7$

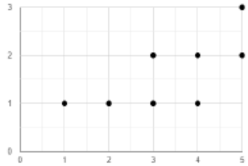


Fig. A5: $f_3(7) = 8$

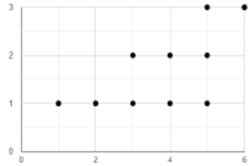


Fig. A6: $f_3(8) = 10$

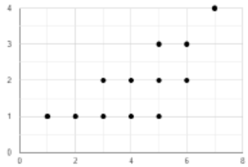


Fig. A7: $f_3(9) = 12$

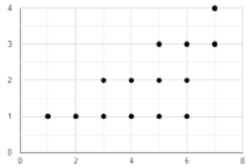


Fig. A8: $f_3(10) = 14$

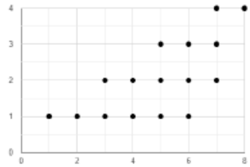


Fig. A9: $f_3(11) = 16$

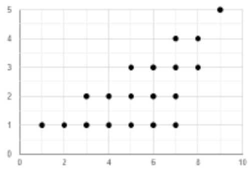


Fig. A10: $f_3(12) = 19$

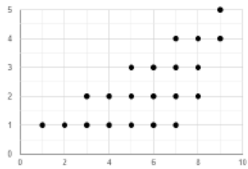


Fig. A11: $f_3(13) = 21$

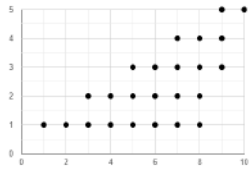


Fig. A12: $f_3(14) = 24$

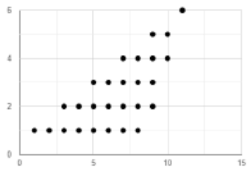


Fig. A13: $f_3(15) = 27$

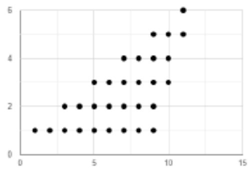


Fig. A14: $f_3(16) = 30$

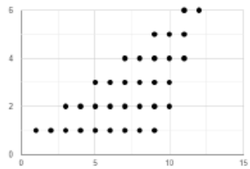
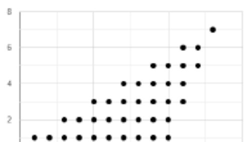
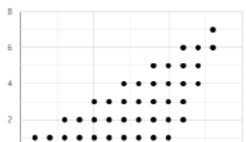
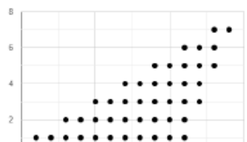


Fig. A15: $f_3(17) = 33$







Construction of identifying and real M-estimators in general statistical model with filtration

T. Toronjadze[1,2]

[1]Georgian American University, Business School, 10 Merab Aleksidze Str., 0160, Tbilisi, Georgia;
[2]A. Razmadze Mathematical Institute of I. Javakhishvili Tbilisi State University, 2 Merab Alekside II Lane, 0193 Tbilisi, Georgia

Abstract

General statistical model with filtration is considered. Identifying and real M-estimators are constructed. Namely, consistent, linear, asymptotically normal estimators are founded, which are basic class of estimators in robust statistics.

Key words and phrases: general statistical model with filtration, M-estimators, robust statistics.

MSC 2010: 62F12, 62F35.

A key part in robust estimation theory play the Huber M-estimators. In general, M-estimators may be viewed as follows.

Consider a sequence of filtered statistical models

$$\mathcal{E} = \left\{(\Omega^n, \mathcal{F}^n, F^n = (\mathcal{F}_t^n),\ 0 \le t \le T,\ (Q_\theta^n,\, \theta \in \Theta \subset R_1))\right\}_{n\ge 1}, \qquad (1)$$

where for each $n \ge 1$ and $\theta \ne \theta'$, the probability measures Q_θ^n and $Q_{\theta'}^n$ are equivalent, $Q_\theta^n \sim Q_{\theta'}^n$, $\mathcal{F}^n = \mathcal{F}_T^n$ and $T > 0$ is a number, σ-algebra $\mathcal{F}_n$ is completed and filtration F^n satisfies the usual conditions w.r.t. Q_θ^n for some, and hence, for each θ.

Let for each $\theta \in \Theta$ and $n \geq 1$ the process $(L_n(\theta,t),\ 0 \leq t \leq T)$ be a local (square integrable) Q_θ^n-martingale.

Denote $L_n(\theta) = L_n(\theta,t)|_{t=T}$ and consider stochastic equation (with respect to parameter θ)

$$L_n(\theta) = L_n(\theta,\omega) = 0, \quad n \geq 1. \tag{2}$$

A sequence $\{T_n(\omega),\ \omega \in \Omega^n\}_{n\geq 1}$ of $\mathcal{F}^n$-measurable roots of these equations $\big($i.e., for each $n \geq 1$, $T_n(\omega)$ is a random variable defined on $(\Omega^n, \mathcal{F}^n)$ with values Θ, and such that

$$L_n(T_n(\omega),\omega) = 0\,\big) \tag{3}$$

is called a generalized M-estimator.

Notice that the equality (3) may be satisfied only asymptotically (in some sense, see, e.g., Theorem 1 below).

The proof of assertions concerning the asymptotic behaviour of M-estimators as solutions of equation (2) is carried out in two steps: firstly, the asymptotic properties are established for the left-hand side of equation (2); secondly, the asymptotic properties of the estimators (considered as implicit functions) are obtained by linearization. In this way one may construct consistent, linear, asymptotically normal estimators, which are asymptotically equivalent of M-estimators (see, e.g., (15) below). Class of such estimators is a basic class of estimators in robust estimation theory (see, e.g., [1, 2, 3]).

1 Local limiting behaviour of roots

Given a sequence of statistical models (1), and let $\{c_n(\theta)\}_{n\geq 1}$, $c_n(\theta) > 0$, $\theta \in \Theta$ be a normalizing deterministic sequence.

Consider the sequence of random variables $\{L_n(\theta)\}_{n\geq 1} = \{L_n(\theta,\omega),\ \omega \in \Omega^n\}_{n\geq 1}$ depending on the parameter $\theta \in \Theta$.

Remark 1. We shall use the following abbreviation

$$Q_\theta^n\text{-}\lim_{n\to\infty} \xi_n = K,$$

where $\xi = \{\xi_n\}_{n\geq 1}$ is a sequence of random variables defined for each n on Ω^n and K is a real number, if $\forall \rho > 0$,

$$\lim_{n\to\infty} Q_\theta^n\{\omega \in \Omega^n : |\xi_n(\omega) - K| > \rho\} = 0.$$

Theorem 1. *Let the following conditions hold:*

a) *for each* $\theta \in \Theta$, $\lim\limits_{n\to\infty} c_n(\theta) = 0$;

b) *for each* $n \geq 1$, *the mapping* $\theta \rightsquigarrow L_n(\theta)$ *is continuously differentiable in* θ Q_θ^n*-a.s.,* $(\dot{L}_n(\theta) := \frac{\partial}{\partial\theta} L_n(\theta))$;

c) *for each* $\theta \in \Theta$, *there exists a function* $\Delta_Q(\theta, y)$, $\theta, y \in \Theta$, *such that*

$$Q_\theta^n\text{-}\lim_{n\to\infty} c_n^2(\theta) L_n(y) = \Delta_Q(\theta, y) \tag{4}$$

and the equation

$$\Delta_Q(\theta, y) = 0$$

with respect to the variable y *has the unique solution* $\theta^* = b^Q(\theta)$;

d) $Q_\theta^n\text{-}\lim\limits_{n\to\infty} c_n^2(\theta)\dot{L}_n(\theta^*) = -\gamma_Q(\theta)$, *where* $\gamma_Q(\theta)$ *is a positive number for each* $\theta \in \Theta$;

e) $\lim\limits_{r\to 0} \limsup\limits_{n\to\infty} Q_\theta^n\{\sup_{\{y:|y-\theta^*|\leq r\}} c_n^2(\theta)|\dot{L}_n(y) - \dot{L}_n(\theta^*)| > \rho\} = 0$ *for each* $\rho > 0$.

Then for each $\theta \in \Theta$ *there exists a sequence of random variables* $T = \{T_n\}_{n\geq 1}$ *taking the values in* Θ *such that*

I. $\lim\limits_{n\to\infty} Q_\theta^n\{L_n(T_n) = 0\} = 1$;

II. $Q_\theta^n\text{-}\lim\limits_{n\to\infty} T_n = \theta^*$;

III. *if* $\{\widetilde{T}_n\}_{n\geq 1}$ *is another sequence with properties I and II, then*

$$\lim_{n\to\infty} Q_\theta^n\{T_n = \widetilde{T}_n\} = 1.$$

If, in addition,

f) *the sequence of distributions* $\{\mathcal{L}\{c_n(\theta)L_n(\theta^*) \mid Q_\theta^n\}\}_{n\geq 1}$ *weakly converges to a certain distribution* Φ,

then

IV. (i) $\mathcal{L}\{\gamma_Q(\theta)c_n^{-1}(\theta)(T_n - \theta^*) \mid Q_\theta^n\} \xrightarrow{w} \Phi$,

$$\text{(ii)}\ \ c_n^{-1}(\theta)(T_n - \theta^*) = \frac{c_n^{-1}(\theta)L_n(\theta^*)}{\gamma_Q(\theta)} + R_n(\theta), \quad R_n(\theta) \xrightarrow{Q_\theta^n} 0.$$

Proof. 1. By the Taylor formula we have

$$L_n(y) = L_n(\theta^*) + \dot{L}_n(\theta^*)(y - \theta^*) + [\dot{L}_n(\bar{\theta}) - \dot{L}_n(\theta^*)](y - \theta^*),$$

where $\bar{\theta} = \theta^* + \alpha(\theta^*)(y - \theta^*)$, $\alpha(\theta^*) \in [0, 1]$ and the point $\bar{\theta}$ is chosen so that $\bar{\theta} \in \mathcal{F}^n$ ($\xi \in \mathcal{F}$ means that r.v. ξ is $\mathcal{F}$-measurable).

From this we get

$$c_n^2(\theta)L_n(y) = c_n^2(\theta)L_n(\theta^*) - \gamma_Q(\theta)(y - \theta^*) + \varepsilon_n(\bar{\theta}, \theta^*)(y - \theta^*), \tag{5}$$

where $\varepsilon_n(y, \theta^*) \in \mathcal{F}^n$,

$$\varepsilon_n(y, \theta^*) = c_n^2(\theta)[\dot{L}_n(y) - \dot{L}_n(\theta^*)] + [c_n^2(\theta)\dot{L}_n(\theta^*) + \gamma_Q(\theta)], \quad y \in \Theta.$$

Evidently, conditions d) and e) ensure that

$$\lim_{r \to 0} \limsup_{n \to \infty} Q_\theta^n \Big\{ \sup_{\{y : |y - \theta^*| \le r\}} |\varepsilon_n(y, \theta^*)| > \rho \Big\} = 0 \tag{6}$$

for each $\rho > 0$.

2. We now show that there exists a family $\{\Omega_\theta(n, r) : n \ge 1,\ r > 0,\ \theta \in \Theta\}$ with properties

$$\begin{aligned} &1) \quad \Omega_\theta(n, r) \in \mathcal{F}^n, \\ &2) \quad \lim_{r \to 0} \limsup_{n \to \infty} Q_\theta^n\{\Omega_\theta(n, r)\} = 1, \end{aligned}$$

and for any $r > 0$, $n \ge 1$ and $\omega \in \Omega_\theta(n, r)$ the equation

$$L_n(y) = 0$$

has the unique solution T_n in the segment $|y - \theta^*| \le r$.

Expansion (5) yields

$$c_n^2(\theta)L_n(\theta^* + u)u = c_n^2(\theta)L_n(\theta^*)u - u^2\gamma_Q(\theta) + u^2\varepsilon_n(\bar{\theta}, \theta^*). \tag{7}$$

For any $\theta \in \Theta$, $n \ge 1$ and $r > 0$ define

$$\Omega_\theta(n, r) = \Big\{ \omega \in \Omega^n : \ |c_n^2(\theta)L_n(\theta^*)| \le \frac{\gamma_Q(\theta)r}{2}, \quad \sup_{\{y : |y - \theta^*| \le r\}} |\varepsilon_n(y, \theta^*)| < \frac{\gamma_Q(\theta)}{2} \Big\}.$$

Obviously, $\Omega_\theta(n,r) \in \mathcal{F}^n$. Hence, if $\omega \in \Omega_\theta(r,n)$, then from equality (7) we get $L_n(\theta^* + u)u < 0$ for $|u| = r$.

Since the mapping $u \rightsquigarrow L_n(\theta^* + u)$ is continuous with respect to u, the equation $L_n(\theta^* + u) = 0$ for $|u| \le r$ has at least one solution $u_n(\theta^*)$ with $|u_n(\theta^*)| \le r$.

It can be easily seen that if $\omega \in \Omega_\theta(n,r)$ and $|u| \le r$, then $\dot{L}_n(\theta^* + u) < 0$.

On the other hand, for $\omega \in \Omega_\theta(n,r)$ and $|u| \le r$,

$$\begin{aligned} &L_n(\theta^* + u, \omega) - L_n(\theta^* + u_n(\theta), \omega) \\ &\qquad = \int_0^1 \frac{\partial}{\partial \alpha} \left[L_n((\theta^* + u_n(\theta^*)) + \alpha(u - u_n(\theta^*)), \omega) \right] d\alpha. \end{aligned}$$

Consequently,

$$L_n(\theta^* + u, \omega) = \int_0^1 \dot{L}(\theta^* + u_n(\theta^*) + \alpha(u - u_n(\theta^*)), \omega)(u - u_n(\theta^*))\, d\alpha$$

and

$$\begin{aligned} &L_n(\theta^* + u, \omega)(u - u_n(\theta^*)) \\ &\qquad = \int_0^1 \dot{L}(\theta^* + u_n(\theta^*) + \alpha(u - u_n(\theta^*)), \omega)(u - u_n(\theta^*))^2\, d\alpha < 0, \end{aligned}$$

provided $u \neq u_n(\theta^*)$. Hence $L_n(\theta^* + u, \omega) \neq 0$ for $|u| \le r$, $u \neq u_n(\theta^*)$. By the construction of the set $\Omega_\theta(n,r)$ and due to conditions c), d) and e) it is easily seen that 2) is true as well.

3. Now we construct the sequence $T = \{T_n\}_{n \ge 1}$ with properties I, II and III. Define

$$\Omega_n^\theta := \bigcup_{k>0} \Omega_\theta(n, k^{-1}).$$

Evidently, $\Omega_n^\theta \in \mathcal{F}^n$. Let $\omega \in \Omega_n^\theta$. Then from the previous statement it follows that there exists a number $k(\omega) > 0$ such that the equation $L_n(y) = 0$ has the unique solution $\widetilde{T}_n(\omega)$ in the segment $|y - \theta^*| \le (k(\omega))^{-1}$ with the mapping $\omega \rightsquigarrow \widetilde{T}_n(\omega)$ which is $\Omega_n^\theta \cap \mathcal{F}^n$-measurable.

Put

$$T_n(\omega) = \begin{cases} \widetilde{T}_n(\omega) & \text{if } \omega \in \Omega_n^\theta, \\ \theta_0 & \text{if } \omega \neq \Omega_n^\theta, \end{cases}$$

where θ_0 is a point in Θ.

It is easily seen that, by construction, T_n possesses properties I, II and III.

4. Finally, we prove assertion IV. By expansion (5), we have

$$
\begin{aligned}
&|c_n(\theta)L_n(T_n) - c_n(\theta)L_n(\theta^*) - \gamma_Q(\theta)c_n^{-1}(\theta)(T_n - \theta^*)| \\
&\qquad \le |\varepsilon_n(\bar{T}, \theta^*)\gamma_Q^{-1}(\theta)|\,|\gamma_Q(\theta)c_n^{-1}(\theta)(T_n - \theta^*)| \qquad (8)
\end{aligned}
$$

and $\limsup_{n\to\infty} Q_\theta^n\{|\varepsilon_n(\bar{T}_n, \theta^*)| \ge \rho\} = 0$, $\forall \rho > 0$, which follows directly from the relation

$$
\{|\bar{T}_n - \theta^*| \le r\} \cap \Big\{ \sup_{\{y:|y-\theta^*|\le r\}} |\varepsilon_n(y, \theta^*)| < \rho \Big\} \subset \{|\varepsilon_n(\bar{T}_n, \theta^*)| < \rho\}.
$$

Denote $X_n := c_n(\theta)(L_n(T_n) - L_n(\theta^*))$, $Y_n := \gamma_Q(\theta)c_n^{-1}(\theta)(T_n - \theta^*)$ and $Z_n := |\varepsilon_n(\bar{T}_n, \theta^*)\gamma_Q^{-1}|$. Then inequality (8) takes the form

$$
|X_n - Y_n| \le Z_n|Y_n|.
$$

It is well-known that if X_n converges weakly to X ($X_n \xrightarrow{w} X$) and $Z_n \xrightarrow{P} 0$, then $Y_n \xrightarrow{w} X$. Thus we get

$$
\lim_{n\to\infty} \mathcal{L}\{\gamma_Q(\theta)c_n^{-1}(\theta)(T_n - \theta^*) \mid Q_\theta^n\} = \lim_{n\to\infty} \mathcal{L}\{c_n(\theta)L_n(\theta^*) \mid Q_\theta^n\}.
$$

Assertion (i) is proved. The proof of assertion (ii) easily follows from (i) and inequality (8). □

2 Global limiting behaviour of roots

We use the objects introduced in the previous section.

Assume $\Theta = [a, b]$. Furthermore, for convenience, put $a = -\infty$ and $b = +\infty$.

For every θ we consider the set

$$
\begin{aligned}
S_\theta = \Big\{\widehat{T} = \{\widehat{T}_n\}_{n\ge1} :{}& \text{for each } n \ge 1,\ \widehat{T}_n \in \mathcal{F}^n \ \text{ and} \\
& Q_\theta^n\text{-}\lim_{n\to\infty} c_n^2(\theta)L_n(\widehat{T}_n) = 0\Big\}.
\end{aligned}
$$

Theorem 2. *Let the following conditions* $(\sup c)$ *hold:*

$(\sup c)_1$ *the function* $\Delta_Q(\theta, y)$ *is* y*-continuous for every* θ*;*

$(\sup c)_2$ *for any* K*,* $0 < K < \infty$*, and* $\rho > 0$*,*

$$\lim_{n\to\infty} Q_\theta^n\Big\{\sup_{|y|\le K} |c_n^2(\theta)L_n(y) - \Delta_Q(\theta, y)| > \rho\Big\} = 0.$$

Then

I. *The following alternative holds: if* $\widehat{T} \in S_\theta$*, then either*

$$Q_\theta^n\text{-}\lim_{n\to\infty} \widehat{T}_n = \theta^* = b^Q(\theta), \tag{9}$$

or

$$\overline{\lim_{n\to\infty}}\, Q_\theta^n\{|\widehat{T}_n| > K\} > 0 \tag{10}$$

for any K*,* $0 < K < \infty$*.*

II. *If, in addition, the condition*

(c^+)
$$\lim_{|y|\to\infty} |\Delta_Q(\theta, y)| = K(\theta) > 0$$

holds and

$$\lim_{n\to\infty} Q_\theta^n\Big\{\sup_{-\infty<y<+\infty} |c_n^2(\theta)L_n(y) - \Delta_Q(\theta, y)| > \rho\Big\} = 0$$

for any $\rho > 0$*, then* (9) *is valid.*

Proof. Let $\widehat{T} = \{\widehat{T}_n\}_{n\ge1} \in S_\theta$ and suppose that inequality (10) is not satisfied. Then there is a number $K_0 > 0$ such that

$$\lim_{n\to\infty} Q_\theta^n\{|\widehat{T}_n| > K_0\} = 0.$$

Therefore,

$$\begin{aligned}
&Q_\theta^n\big\{|c_n^2(\theta)L_n(\widehat{T}_n) - \Delta_Q(\theta, \widehat{T}_n)| > \rho\big\} \\
&\quad \le Q_\theta^n\big\{|\widehat{T}|_n > K_0\big\} + Q_\theta^n\Big\{|c_n^2(\theta)L_n(\widehat{T}_n) - \Delta_Q(\theta, \widehat{T}_n)| > \rho,\ \ |\widehat{T}_n| \le K_0\Big\} \\
&\le Q_\theta^n\big\{|\widehat{T}|_n > K_0\big\} + Q_\theta^n\Big\{\sup_{|y|\le K_0} |c_n^2(\theta)L_n(y) - \Delta(\theta, y)| > \rho\Big\} \to 0 \ \text{ as } \ n \to \infty.
\end{aligned}$$

On the other hand,

$$Q_\theta^n\text{-}\lim_{n\to\infty} c_n^2(\theta)L_n(\widehat{T}_n) = 0$$

and hence,

$$Q_\theta^n\text{-}\lim_{n\to\infty} \Delta_Q(\theta, \widehat{T}_n) = 0. \tag{11}$$

Assume now that equality (9) fails too. Then one can choose $\varepsilon > 0$ such that

$$\overline{\lim_{n\to\infty}}\, Q_\theta^n\big\{|\widehat{T}_n - b^Q(\theta)| > \varepsilon\big\} > 0.$$

By the condition $(\sup c)_1$,

$$\Delta(\varepsilon) = \inf_{\{y:|y-b^Q(\theta)|>\varepsilon,\, |y|\le K_0\}} |\Delta_Q(\theta, y)| > 0,$$

whence

$$\begin{aligned}
&\overline{\lim_{n\to\infty}}\, Q_\theta^n\big\{|\Delta_Q(\theta, \widehat{T}_n)| > \Delta(\varepsilon)\big\}\\
&\qquad \ge \overline{\lim_{n\to\infty}}\, Q_\theta^n\big\{|\Delta_Q(\theta, \widehat{T}_n)| > \Delta(\varepsilon),\ \ |\widehat{T}_n| \le K_0\big\}\\
&\qquad \ge \overline{\lim_{n\to\infty}}\, Q_\theta^n\big\{|\widehat{T}_n - b^Q(\theta)| > \varepsilon,\ \ |\widehat{T}_n| \le K_0\big\} > 0,
\end{aligned}$$

which contradicts equality (11).

In order to prove the second assertion of theorem, it is sufficient to note that under the condition (c^+)

$$\inf_{\{y:|y-b^Q(\theta)|\ge\varepsilon\}} |\Delta_Q(\theta, y) > 0$$

and to repeat the previous arguments. □

Suppose that the conditions of Theorem 1 are satisfied.

For every $n \ge 1$, consider the set

$$A_n = \{\omega \in \Omega^n : \text{ the equation } L_n(y, \omega) = 0 \text{ has at least one solution}\}.$$

Note that $A_n \in \mathcal{F}^n$. Indeed, recall that the σ-algebra $\mathcal{F}^n$ is complete, $L_n(y, \cdot) \in \mathcal{F}^n$ for each fixed y and $L_n(\cdot, \omega)$ is a.s. continuous. Hence, the mapping $(y, \omega) \rightsquigarrow L_n(y, \omega)$ is measurable and $B_n := \{(y, \omega) : L_n(y, \omega) = 0\} \in \mathcal{B}(R_1) \times \mathcal{F}^n$. But $A_n = \Pi_{\Omega^n}(B_n)$, where $\Pi_{\Omega^n}(\cdot)$ is a projection operator. Thus $A_n \in \mathcal{F}^n$.

Evidently, for any θ, we have $\Omega_n^\theta \subset A_n$, where the set Ω_n^θ is defined in item 3 of the proof of Theorem 1.

Since under the conditions of Theorem 1, $Q_\theta^n\{\Omega_n^\theta\} \to 1$, for any θ we have

$$\lim_{n\to\infty} Q_\theta^n\{A_n\} = 1.$$

For each $n \geq 1$, introduce the sets:

$$S_n = \{\widetilde{T}_n :\ \widetilde{T}_n \text{ is } \mathcal{F}^n\text{-measurable; } L_n(\widetilde{T}_n) = 0 \text{ if } \omega \in A_n;\ \widetilde{T}_n = \theta_0 \text{ if } \omega \notin A_n\},$$

where θ_0 is a real number.

Now, put the set of estimators

$$S_{sol} = \{\widetilde{T} = \big\{\widetilde{T}_n\}_{n\geq 1} : \forall n \geq 1,\ \widetilde{T}_n \in S_n\big\}.$$

Corollary 1. *If along with the conditions of Theorem 1 the conditions* $(\sup c)$ *are satisfied for any θ, then there exists an estimator $T^* = \{T_n^*\}_{n\geq 1} \in S_{sol}$ such that*

$$Q_\theta^n\text{-}\lim_{n\to\infty} T_n^* = b^Q(\theta) \tag{12}$$

for any θ.

If, moreover, for any θ the condition (c^+) *is satisfied, then any estimator $\widetilde{T} \in S_{sol}$ has property* (12).

Proof. It is sufficient to construct an estimator $T^* = \{T_n^*\}_{n\geq 1}$ for which (10) fails for each θ.

For any $n \geq 1$ and $\varepsilon > 0$, there exists $T_n^* \in S_n$ such that

$$|T_n^*| \leq \operatorname*{ess\,inf}_{\widetilde{T}_n \in S_n} |\widetilde{T}_n| + \varepsilon.$$

By virtue of Theorem 1, for any θ there exists a sequence $\widehat{T}(\theta) = \{\widehat{T}_n(\theta)\}_{n\geq 1}$ such that

$$\lim_{n\to\infty} Q_\theta^n\{L_n(\widehat{T}_n(\theta)) = 0\} = 1 \tag{13}$$

and

$$Q_\theta^n\text{-}\lim_{n\to\infty} \widehat{T}_n(\theta) = b^Q(\theta). \tag{14}$$

Thus, we have

$$\begin{aligned}
\overline{\lim_{n\to\infty}}\, Q_\theta^n\big\{|T_n^*| > K\big\} &\leq \overline{\lim_{n\to\infty}}\, Q_\theta^n\big\{|T_n^*| > K,\ L_n(\widehat{T}_n(\theta)) \neq 0\big\} \\
&\quad + \overline{\lim_{n\to\infty}}\, Q_\theta^n\big\{|T_n^*| > K,\ L_n(\widehat{T}_n(\theta)) = 0\big\} \\
&\leq \overline{\lim_{n\to\infty}}\, Q_\theta^n\big\{L_n(\widehat{T}_n(\theta)) \neq 0\big\} + \overline{\lim_{n\to\infty}}\, Q_\theta^n\big\{|\widehat{T}_n(\theta)| + \varepsilon > K\big\}.
\end{aligned}$$

The first and the second terms on the right-hand side converge to zero by virtue of equalities (13) and (14). $\square$

Remark 2. If the conditions of Corollary 1 are satisfied, then by virtue of Theorem 1, IV (ii), there exists an estimator $T = \{T_n\}_{n \geq 1}$ such that

$$T_n = \theta^* + \frac{L_n(\theta^*)}{\gamma_Q(\theta)} + R_n(\theta), \tag{15}$$
$$c_n^{-1}(\theta) R_n(\theta) \xrightarrow{Q_\theta^n} 0.$$

If $\theta^* = b^Q(\theta) = \theta$ and the distribution Φ from Theorem 1, f), is Gaussian, then we obtain a consistent, linear, asymptotically normal estimator.

References

[1] P. J. Huber, *Robust Statistics*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1981.

[2] P. J. Huber, E. M. Ronchetti, *Robust Statistics*. Second edition. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, 2009.

[3] R. A. Maronna, R. D. Martin, V. J. Yohai, M. Salibián-Barrera, *Robust Statistics. Theory and Methods (with R)*. Second edition of [MR2238141]. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, 2019.

ON A GENERALIZATION OF KHINCHIN'S THEOREM

V. BERIKASHVILI, G. GIORGOBIANI, V. KVARATSKHELIA

ABSTRACT. A generalization of Khinchin's theorem for weakly correlated random elements with values in Banach spaces $l_p, 1 \le p < \infty$ is presented without proof.

The purpose of this paper is to generalize the following Khinchin's theorem, which was published in 1928 in the journal of the French Academy of Sciences [1]. The concepts and background information about probability distributions in infinite-dimensional spaces, necessary for further discussion, can be found in [2].

Let $\xi_1, \xi_2, \dots, \xi_n, \dots$ be a sequence of real random variables, defined on the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ with finite mathematical expectations $\mathbb{E}\,\xi_n < \infty$; denote $S_n = \sum_{i=1}^{n} \xi_i, n = 1, 2, \dots$. We say that the given sequence of random variables satisfies the *Law of Large Numbers* (*LLN*), if the sequence $\{\frac{S_n - \mathbb{E}\,S_n}{n}\}$ converges in probability to zero as $n \to \infty$, i.e. for every $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left[\left|\frac{S_n - \mathbb{E}\,S_n}{n}\right| > \varepsilon\right] = 0.$$

Theorem 1. (A.Y. KHINCHIN, 1928). *Let $\xi_1, \xi_2, \dots, \xi_n, \dots$ be a sequence of random variables such that for any positive integer n, ξ_n has a finite mathematical expectation and variance σ_n^2. Furthemore, let g be a nonnegative function, defined on the set of nonnegative integers such that for the correlation coefficients ϱ_{mn} of ξ_m and ξ_n the following inequalities hold*

$$|\varrho_{mn}| \le g(|m-n|), \qquad m, n = 1, 2, \dots.$$

If

$$\lim_{n\to\infty} \frac{1}{n^2} \left(\sum_{i=0}^{n-1} g(i)\right) \left(\sum_{i=1}^{n} \sigma_i^2\right) = 0, \tag{1}$$

then the sequence $\xi_1, \xi_2, \dots, \xi_n, \dots$ satisfies the LLN.

In this paper [1] A. Khinchin introduced the well-established term "Strong Law of Large Numbers" (*SLLN*) and proved that the theorem formulated above provides a sufficient condition for the fulfillment of the *SLLN* if condition (1) is replaced by the condition

$$\lim_{n\to\infty} \frac{1}{n^{2-\delta}} \left(\sum_{i=0}^{n-1} g(i)\right) \left(\sum_{i=1}^{n} \sigma_i^2\right) = 0$$

for some $\delta > 0$.

Recall some general notions. Let X be a separable Banach space with a norm $\|\cdot\|$, X^* be its dual, $\langle x^*, x\rangle$ be a value of the functional $x^* \in X^*$ at the point $x \in X$, $(\Omega, \mathfrak{F}, \mathbb{P})$ be a fixed probability space. Denote by $\mathfrak{B}(X)$ the Borel σ-algebra in X. A map $\xi : \Omega \to X$ is called *a random element* with values in X if $\xi^{-1}\{\mathfrak{B}(X)\} \subset \mathfrak{F}$.

1991 *Mathematics Subject Classification.* 60B12; 60B11.

Key words and phrases. Law of large numbers, l_p spaces, covariance operator, cross-covariance operator, correlation coefficient.

It is said that a random element ξ with values in X has *a weak p-order*, $p > 0$, if $\mathbb{E}\,|\langle x^*, \xi\rangle|^p < \infty$ for every $x^* \in X^*$. If a random element ξ has a weak p-order, $p > 1$, then the *expectation* $\mathbb{E}\,\xi$ exists and is defined as the *Pettis integral* of the random element ξ. Without loss of generality we assume that all random elements considered below are centered (that is, $\mathbb{E}\,\xi{=}0$). For a random element ξ with the weak second order *covariance operator* $R_\xi : X^* \to X$ is defined as follows:

$$\langle x^*, R_\xi x^*\rangle = \mathbb{E}\,\langle x^*, \xi\rangle^2, \qquad x^* \in X^*.$$

It is easy to see that R_ξ is a nonnegative, symmetric and linear continuous operator. For any symmetric nonnegative operator $R : X^* \to X$ there is a Hilbert space H and a linear continuous operator $A : X^* \to H$ such that $R = A^*A$; the operator A is uniquely determined up to isometry (see [2], Factorization Lemma, p. 123).

Random element $\xi : \Omega \to X$ is called *Gaussian* if $\langle x^*, \xi\rangle$ is a Gaussian random variable for any $x^* \in X^*$. We say that an operator $R : X^* \to X$ is a *Gaussian covariance* if there exists a Gaussian random element with values in X such that its covariance operator coincides with R.

Let $X = H$ be a Hilbert space with the inner product $(\cdot,\cdot)_H$. An operator $T : H \to H$ is called *nuclear* if it admits the representation

$$Th = \sum_{i=1}^{\infty}(a_i, h)_H b_i \qquad h \in H,$$

and for some sequences $\{a_i\}$ and $\{b_i\}$ in H with $\sum\limits_{i=1}^{\infty} \|a_i\|_H\|b_i\|_H < \infty$.

Let $\{\varphi_k\}$ be an orthonormal basis in H. Then for the nuclear operator $T : H \to H$ the series

$$\sum_{k=1}^{\infty}(T\varphi_k, \varphi_k)_H \tag{2}$$

converges, the sum (2) does not depend on the choice of the orthonormal basis and is called the *trace* ($tr(T)$) of the operator T. If a random element ξ with values in a Hilbert space has a strong second order ($\mathbb{E}\,\|\xi\|_H^2 < \infty$) and R_ξ is its covariance operator, then it is easy to see that $\mathbb{E}\,\|\xi\|_H^2 = tr(R_\xi)$.

Let ξ and η be random elements of weak second order with values in a Banach space X (recall that $\mathbb{E}\,\xi = \mathbb{E}\,\eta = 0$). *Cross-covariance operator* $R_{\xi\eta} : X^* \to X$ of ξ and η is defined by the equality:

$$\langle x^*, R_{\xi\eta}y^*\rangle = \mathbb{E}\,\langle x^*, \xi\rangle\langle y^*, \eta\rangle, \qquad x^*, y^* \in X^*.$$

It is known that $R_{\xi\eta}$ admits the factorization [3]:

$$R_{\xi\eta} = A_\xi^* V_{\xi\eta} A_\eta, \tag{3}$$

where A_ξ (resp. A_η) is a continuous linear operator from X^* to some Hilbert space H_ξ (resp. H_η) such that $R_\xi = A_\xi^* A_\xi$ (resp. $R_\eta = A_\eta^* A_\eta$), the set $A_\xi(X^*)$ (resp. $A_\eta(X^*)$) is dense in H_ξ (resp. in H_η), and $V_{\xi\eta} : H_\eta \to H_\xi$ is a continuous linear operator and for the operator norm we have $\|V_{\xi\eta}\| \le 1$.

$V_{\xi\eta}$ is called a *correlation coefficient* and as in the one-dimensional case, is a measure of the linear dependence of the random elements [3].

To prove the main result we need the following elementary lemma, which was actually applied in [1].

Lemma 2. *Let $\alpha_i, \beta_{i-1}, \varrho_{ij}$, $i,j = 1,2,\ldots,n$, be the sequences of nonnegative numbers and let*

$$\varrho_{ij} \leq \beta_{|i-j|} \quad \textit{for any} \quad i,j = 1,2,\ldots,n.$$

Then

$$\sum_{i,j=1}^{n} \varrho_{ij}\alpha_i\alpha_j \leq 2\left(\sum_{i=0}^{n-1}\beta_i\right)\left(\sum_{i=1}^{n}\alpha_i^2\right).$$

Consider the Banach space of all p-absolutely convergent sequences of real numbers $l_p, 1 \leq p < \infty$, with the norm $\|\cdot\|_{l_p}$. As we know the dual space is $l_p^* = l_q$, $pq = p+q$, when $1 < p < \infty$, and $l_1^* = l_\infty$.

Let ξ be a random element in $l_p, 1 \leq p < \infty$, with the covariance operator R_ξ. Let $e_k = (0,\ldots,\overset{k}{1},0,\ldots)$, $k = 1,2,\ldots$, be a sequence of unit vectors in the dual space l_p^*. Recall that R_ξ is a Gaussian covariance operator if and only if (see [2], Theorem 5.6, p. 261)

$$\sum_{k=1}^{\infty}\langle e_k, R_\xi e_k\rangle^{p/2} < \infty. \tag{4}$$

Let us state the main result of this paper.

Theorem 3. *Let $\xi_1,\xi_2,\ldots,\xi_n,\ldots$ be a sequence of weak second order random elements with values in $l_p, 1 \leq p < \infty$, and let the covariance operator $R_n \equiv R_{\xi_n} : l_p^* \to l_p$, satisfy the condition*

$$\sigma_n^s \equiv \sum_{k=1}^{\infty}\langle e_k, R_n e_k\rangle^{s/2} < \infty, \quad n = 1,2,\ldots, \tag{5}$$

where $s = \min\{2,p\}$. Let, besides there exists a nonnegative function g, defined on the set of nonnegative integers such that for the correlation coefficient V_{mn} of ξ_m and ξ_n the following inequalities hold

$$\|V_{mn}\| \leq g(|m-n|) \qquad \textit{for any} \qquad m,n = 1,2,\ldots.$$

If

$$\lim_{n\to\infty}\frac{1}{n^2}\left(\sum_{i=0}^{n-1}g(i)\right)\left(\sum_{i=1}^{n}\sigma_i^s\right)^{2/s} = 0, \tag{6}$$

then

$$\lim_{n\to\infty}\mathbb{E}\left\|\frac{1}{n}\sum_{i=1}^{n}\xi_i\right\|_{l_p}^{s} = 0. \tag{7}$$

In particular the sequence $\xi_1,\xi_2,\ldots,\xi_n,\ldots$ satisfies the LLN.

Remark 4. *The complete proof of Theorem 3 is published in* [4].

When $p = 2$ Theorem 3 implies the following statement.

Corollary 5. *Let $\xi_1,\xi_2,\ldots,\xi_n,\ldots$ be a sequence of strong second order random elements with values in a separable Hilbert space and let a function g satisfies the requirements of Theorem 3. If*

$$\lim_{n\to\infty}\frac{1}{n^2}\left(\sum_{i=0}^{n-1}g(i)\right)\left(\sum_{i=1}^{n}tr(R_i)\right) = 0,$$

then the sequence $\xi_1,\xi_2,\ldots,\xi_n,\ldots$ satisfies the LLN.

Corollary 6. *Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be a sequence of weak second order random elements with values in $l_p, 1 \leq p \leq 2$, and let the covariance operators $R_n \equiv R_{\xi_n}$ satisfy the condition*

$$\sigma_n^p \equiv \sum_{k=1}^{\infty} \langle e_k, R_n e_k \rangle^{p/2} < \infty, \quad n = 1, 2, \ldots.$$

If

$$\lim_{n\to\infty} \frac{1}{n} \left(\sum_{i=1}^{n} \sigma_i^p \right)^{2/p} = 0, \tag{8}$$

then the sequence $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ satisfies the LLN.

If the random elements are pairwise independent (or not correlated), then obviously we can assume that $\sum_{i=0}^{n-1} g(i) = 1$ for any positive integer n. Thus Theorem 3 implies the following

Corollary 7. *Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be a sequence of pairwise independent weak second order random elements with values in $l_p, 1 \leq p < \infty$, and let for any positive integer n covariance operators $R_n \equiv R_{\xi_n}$ satisfy (3.2).*

If

$$\lim_{n\to\infty} \frac{1}{n^2} \left(\sum_{i=1}^{n} \sigma_i^s \right)^{2/s} = 0, \quad \textit{where} \quad s = \min\{2, p\},$$

then the sequence $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ satisfies the LLN.

In particular, for the case of a separable Hilbert space we have

Corollary 8. *Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be a sequence of pairwise independent strong second order random elements with values in a separable Hilbert space and let*

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} tr(R_i) = 0.$$

Then the sequence $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ satisfies the LLN.

Naturally the question arises about the validity of the main theorem of the paper in the general Banach space. Does it remain true at least in the case of Banach spaces with an unconditional basis and a finite cotype? The answer to this question is not yet known to us.

References

[1] A.Y. Khinchin. Sur la loi forte des grands nombres. C. R. Acad. Sci. Paris Ser. I Math, (1928), **186, 285**, p. 285–287.

[2] Vakhania N.N., Tarieladze V.I., Chobanyan S.A. Probability Distributions on Banach Spaces. D. Reidel Publishing Company, Dordrecht, (1987), 506 p.

[3] Kvaratskhelia V. The analogue of the coefficient of correlation in Banach Spaces. Bull. Georgian Acad. Sci., (2000), **161**, 3, p. 377–379.

[4] Berikashvili V., Giorgobiani G., Kvaratskhelia V. The Law of Large Numbers for weakly Correlated Random Elements in the Spaces $l_p, 1 \leq p < \infty$. Lithuanian Mathematical Journal, 62, 2022, p. 308-314. https://doi.org/10.1007/s10986-022-09564-x

V. Berikashvili, Muskhelishvili Institute of Computational Mathematics, Georgian Technical University, Gr. Peradze str., 4, Tbilisi 0159, Georgia.

E-mail address: `v.berikashvili@gtu.ge`

G. Giorgobiani, Muskhelishvili Institute of Computational Mathematics, Georgian Technical University, Gr. Peradze str., 4, Tbilisi 0159, Georgia.
E-mail address: giorgobiani.g@gtu.ge

V. Kvaratskhelia, Muskhelishvili Institute of Computational Mathematics, Georgian Technical University, Gr. Peradze str., 4, Tbilisi 0159, Georgia.
E-mail address: v.kvaratskhelia@gtu.ge

Real Options Valuation using Machine Learning Methods

Levan Gachechiladze
Business School
Georgian-American University

Abstract

This paper follows work on establishing comprehensive framework for investment projects valuation discussed in [3], [8], [12], and [20]. Previous work is focused on capturing strategic value of investment projects while also incorporating strategic decisions of competitors. Two main methodologies that comprise such valuation framework are Real Options Analysis and Game Theory. In this paper, it is attempted to price real options using Machine Learning (ML) methods. First, selected machine learning models are trained to predict option prices as given by The Black-Scholes formula. Having shown some promise by work discussed in [22] and [23], real-world data has been selected for pricing options and then training machine learning models on them. Finally, various investment projects have been simulated to price option to expand using Cox-Ross-Rubinstein binomial model discussed in [4] and then train machine learning models to predict it. This, in turn, has potential to incorporate market competition implicitly in the value of the strategic option during training process. Hence, machine learning approach can become real options pricing method that is valid not only for monopolistic markets. With this aim, section 1 of the paper gives brief introduction of option pricing methods, section 2 uses Nasdaq Futures historical prices for training ML models to price financial options, and section 3 uses simulated investment projects for training ML models and pricing options to expand. Complete code is available at github.com/leongache/Real-Options-Valuation-using-Machine-Learning-Methods.

1. Overview of classical option pricing methods

This section briefly defines most commonly used methods for option valuation. The purpose of this section is to identify what makes each method usable in the first place and then what makes it less attractive in some practical use cases.

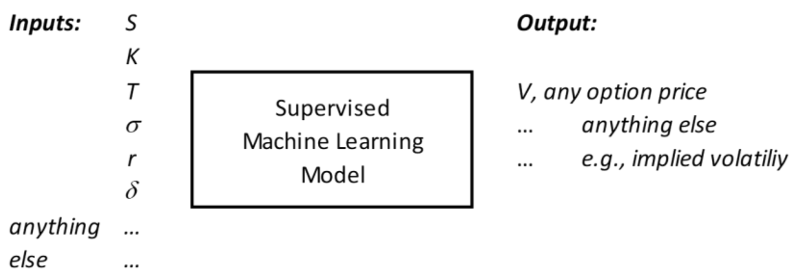
Let's start with The Black-Scholes model, also known as the Black-Scholes-Merton model. Most attractive feature of the model may be its closed form solution, it's simplicity despite complicated mathematics behind it. On the other hand, the Black-Scholes model makes certain assumptions that in certain market conditions may result in prices that deviate from real-world results. E.g., constant risk-free rate and volatility of underlying asset become less evident in extreme/turbulent markets with unpredictable high volatility. Though, improvements of the Black-Scholes model that account for some of its disadvantages exist, its unrealistic assumptions make it hard to use the plain formula for accurate option pricing in all market conditions.

Most flexible option pricing model that can improve on many of the Black-Scholes model limitations is Binomial Option Pricing model. It's commonly used to price American-style options that can be exercised before expiration and have flexibility to price options with any payoff formula as well as incorporate variable inputs for risk-free rate, volatility, etc. in dynamic fashion. Although, it's very hard to predict what those inputs will be equal to in future dates. Even if accurately predicted, incorporating variable inputs increases models' complexity to the point when it may become hard to formulate.

Lastly, commonly used option pricing method is Monte Carlo simulation where numerous random paths for the price of an underlying asset are generated, each having an associated payoff. Then present value of payoffs is computed, and their average becomes an option price that values in all simulated scenarios. Just like Binomial Option Pricing model, this method can incorporate any option payoff and dynamically introduced inputs. Moreover, this method is not restricted to a single or any distribution of underlying asset unlike models with closed-form solutions as given by the Black-Scholes. All that gives Monte Carlo simulation substantial number of use-case in real-life applications. Main disadvantage of the method stays to be heavy computational load as it requires a large number of simulations to improve average accuracy.

Next section in this paper introduces more recent approach to option pricing using artificial intelligence, mainly, machine learning (ML) methods. Using same or more number of inputs as in classical options pricing methodologies ML methods can be trained from both simulated and historical data to "learn" either observed option price or theoretical one given by option pricing method of our choice. Success of ML model will depend on quality of training data and its properties for generalization among other things. In case of creating successful ML model that accurately predicts option prices on out-of-sample data, one can conclude that disadvantages of classical option pricing models will no longer be a concern for both academia and practitioners. At its simplest form, machine learning approach is illustrated in figure below.

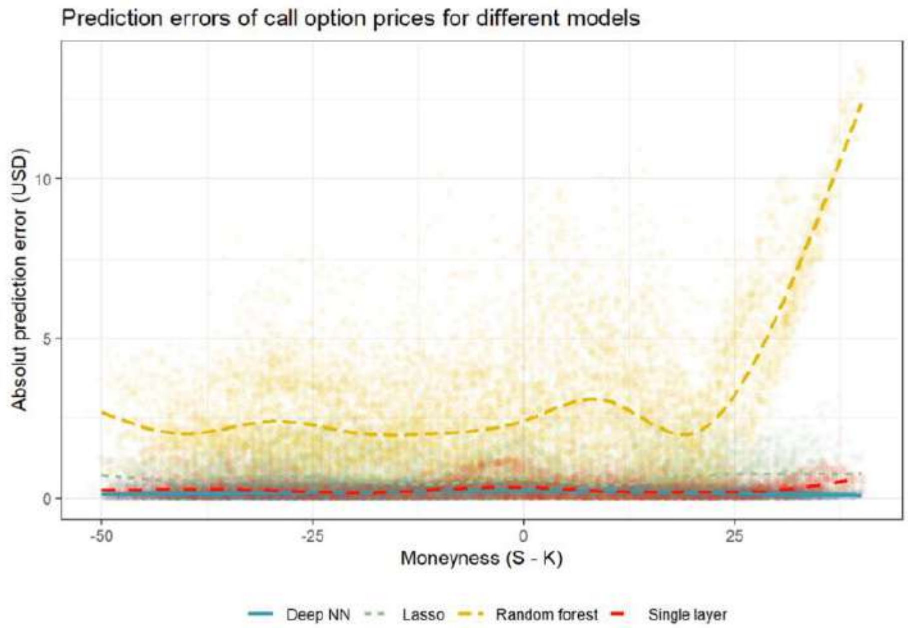
Figure 1 input-output structure of supervise machine learning models



2. Training ML models to learn The Black-Scholes formula

This section carries an experiment of training most popular Machine Learning models to predict call and put option prices as given by The Black-Scholes formula. Building up on work in [22] this example uses real-world prices of continuous Nasdaq Futures contracts publicly available on yahoo finance webpage. Figure 2 shows results from [22] where more than 1.5 million random parameter constellations were used to simulate options prices and train models on them. Figure clearly identifies that all but one ML model under consideration seem to be able to price call options with different moneyness levels.

Figure 2 Prediction error of call options prices for different ML models, source in ref. [22]



Above figure suggests that ML models can be used to price options, and in this section similar experiment is carried out now using actual data instead of simulated one for illustrative purposes.

Using daily prices for Nasdaq futures starting from earliest available date as of 19 Sep 2001, at-the-money call and put prices are computed using following input parameters:

- annual standard deviation of continuously compounded returns as a volatility input,
- annual risk-free rate of 3%, and
- time to expiration of 20 trading days.

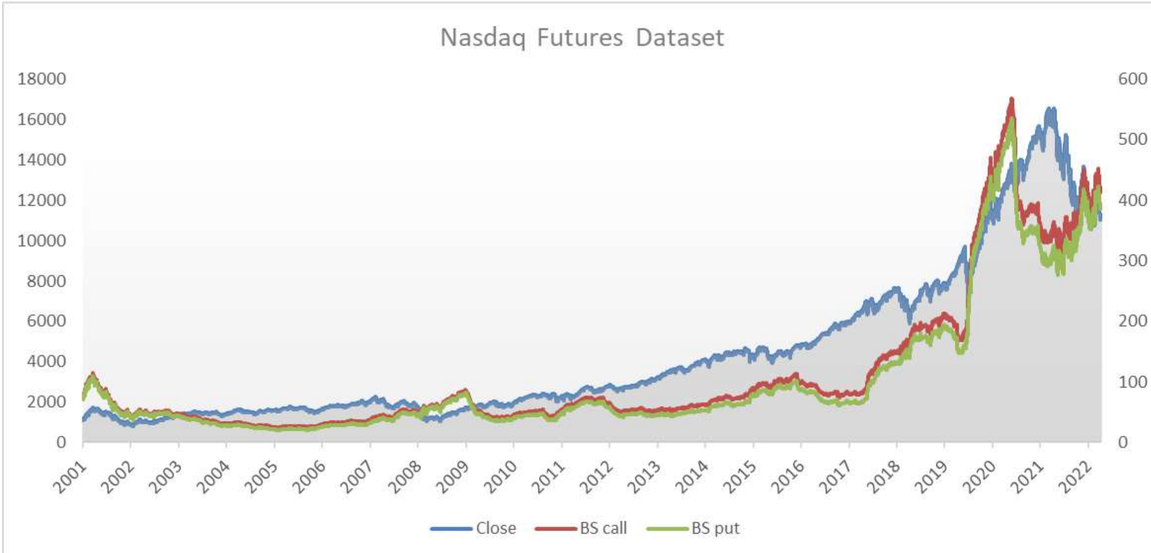
Code snippet below shows functions used to compute option prices for Nasdaq Futures data till most recent date as of time of writing, 23 Dec 2022.

Code Snippet 1 The Black-Scholes call and put prices, source https://www.codearmo.com/python-tutorial/options-trading-black-scholes-model

```
# functions for option prices
N = norm.cdf
def BS_CALL(S, K, T, r, sigma):
    d1 = (np.log(S/K) + (r + sigma**2/2)*T) / (sigma*np.sqrt(T))
    d2 = d1 - sigma * np.sqrt(T)
    return S * N(d1) - K * np.exp(-r*T)* N(d2)


def BS_PUT(S, K, T, r, sigma):
    d1 = (np.log(S/K) + (r + sigma**2/2)*T) / (sigma*np.sqrt(T))
    d2 = d1 - sigma* np.sqrt(T)
    return K*np.exp(-r*T)*N(-d2) - S*N(-d1)
```

Figure below shows complete dataset used for training and testing of ML models; this includes historical data of Nasdaq Futures as well as calculated option prices using formulas in Code Snippet 1.

Figure 3 Nasdaq futures historical prices and option prices by The Black-Scholes (BS) model



Despite high volatility in times series, most recent 5% of complete data was selected as out-of-sample for testing purposes. Out of most common and fundamentally different ML models, the following 4 were selected:

- K-nearest neighbors (KNN)
- Multi-layer Perceptron (MLP)
- Gradient Boosting Trees, specifically, LightGBM implementation of it
- Support Vector Machines (SVM)

In addition, standard scaler was used to standardize training and test data to have standard normal distributions before training any of above models. Moreover, a small sample of parameter distributions were selected in advance to run parameter optimization per each model, separately. For that purposes, Randomized Grid Search algorithm was selected. Specific ML configurations used are shown and described below.

For KNN:

- *leaf_size*: starting from 5, till 100, with steps of 5
- *n-neighbors*: starting from 5, till 100, with steps of 5

For MLP, fixed random state, 500 max iterations, lbfgs[1] solver, tolerance of 1e-8 and:

- *alpha*: [0.01,0.001,0.0001],
- *hidden_layer_sizes*: [(5,5,5,), (5,), (5,5,)]

For LightGBM, subsample and colsample_bytree of .8 and:

- *n_estimators*: starting from 50, till 1000, with steps of 50
- *max-depth*: starting from 3, till 10, with steps of 1

For SVMs:

- *C*: starting from .01, till 1, with steps of .01

[1] lbfgs - an optimizer in the family of quasi-Newton methods that approximates the Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) using a limited amount of computer memory.

- *gamma*: starting from .01, till 1, with steps of .01

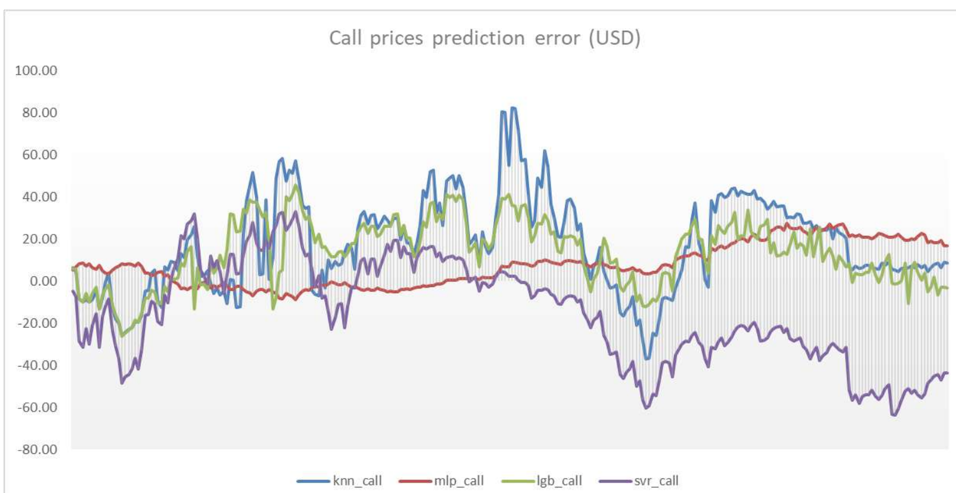
It's worth noting that default configuration for Randomized Grid Search has maximum number of iterations set to 10 and k-fold cross-validation set to 5 folds. Consequently, all models above will get 5-fold cross validation and 10 randomly chosen parameter combinations. Only exception is MLP as it only has 9 possible parameter combinations to search from, in which case exhaustive search will be implemented. It's important to note that there is no right architecture choice for neural networks in general, not so even for such simple networks as MLPs. For that reason, an arbitrary number of nodes and 3 hidden layer variations is chosen without any particular reason. Same is true for distribution ranges of other ML model parameters. Code snippet below summarizes final modules used for an experiment.

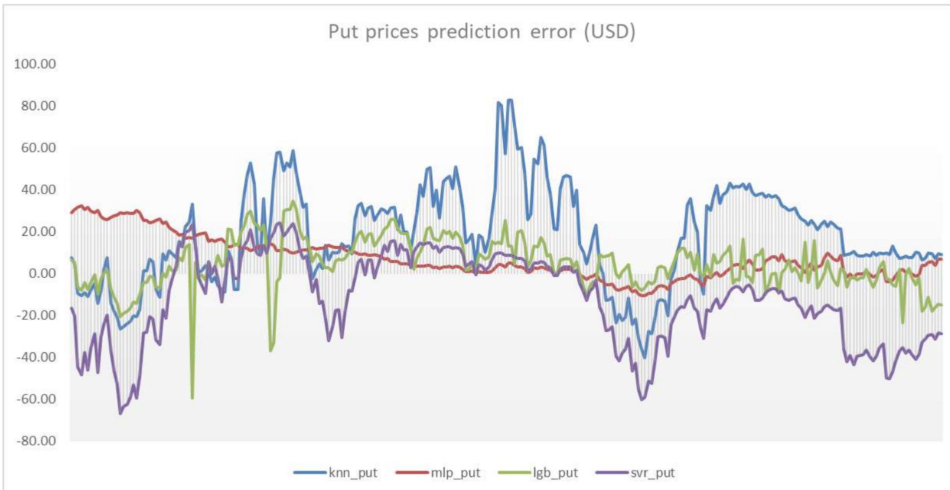
Code Snippet 2 Configuration of machine learning models used for option pricing

```
knn_search = RandomizedSearchCV(KNeighborsRegressor(), {'leaf_size':np.arange(5,100,5),'n_neighbors':np.arange(5,100,5)})

mlp_search = RandomizedSearchCV(MLPRegressor(random_state=0,max_iter=500,solver='lbfgs',tol=1e-8),
                                {'alpha':[0.01,0.001,0.0001],'hidden_layer_sizes':[(5,5,5,),(5,),(5,5,)]})

lgb_search = RandomizedSearchCV(LGBMRegressor(boosting_type='goss',subsample=0.8,colsample_bytree=0.8),
                                {'n_estimators':np.arange(50,1000,50),'max_depth':np.arange(3,10,1)})

svr_search = RandomizedSearchCV(SVR(),{'C':np.arange(.01,1,.01),'gamma':np.arange(.01,1,.01)})
```

Using total of 5380 days of real data, and having selected only 5% for testing purposes, models were trained on 5111 days and tested on next 269 days. Separate training for call and put option prices generated following dollar value errors per model class.

Figure 4 Prediction errors on call and put prices from Nasdaq futures





Considering lack of training data, prediction errors look satisfactory with MLP having smallest errors for both call and put option prices and, hence, can be considered most suitable ML model for the task. Table below summarizes average error and min max ranges for all models individually.

Table 1 Prediction error metrics per ML method

metric	knn_call	mlp_call	lgb_call	svr_call	knn_put	mlp_put	lgb_put	svr_put
average error	22.66	9.61	17.04	23.64	24.06	8.74	9.74	19.33
min error	0.20	0.04	0.00	0.29	0.25	0.00	0.09	0.14
max error	82.39	27.39	45.59	63.58	82.94	32.32	59.41	66.93
max neg error	-37.02	-8.64	-25.99	-63.58	-40.25	-10.54	-59.41	-66.93
max pos error	82.39	27.39	45.59	33.11	82.94	32.32	34.79	24.17

Table above identifies MLP and LGB as better models then KNN and SVR given prediction performance on test dataset under consideration.

Based on promising results from above, next section attempts to train selected ML models to price not financial but Real Options for investment projects valuation.

3. Pricing real options, namely, option to expand using ML methods

Most common examples of Real Options used for strategic investment valuations are:

- Option to Expand
- Option to Contract
- Option to Defer
- Option to Abandon
- Option to Choose
- Option to Switch Resources

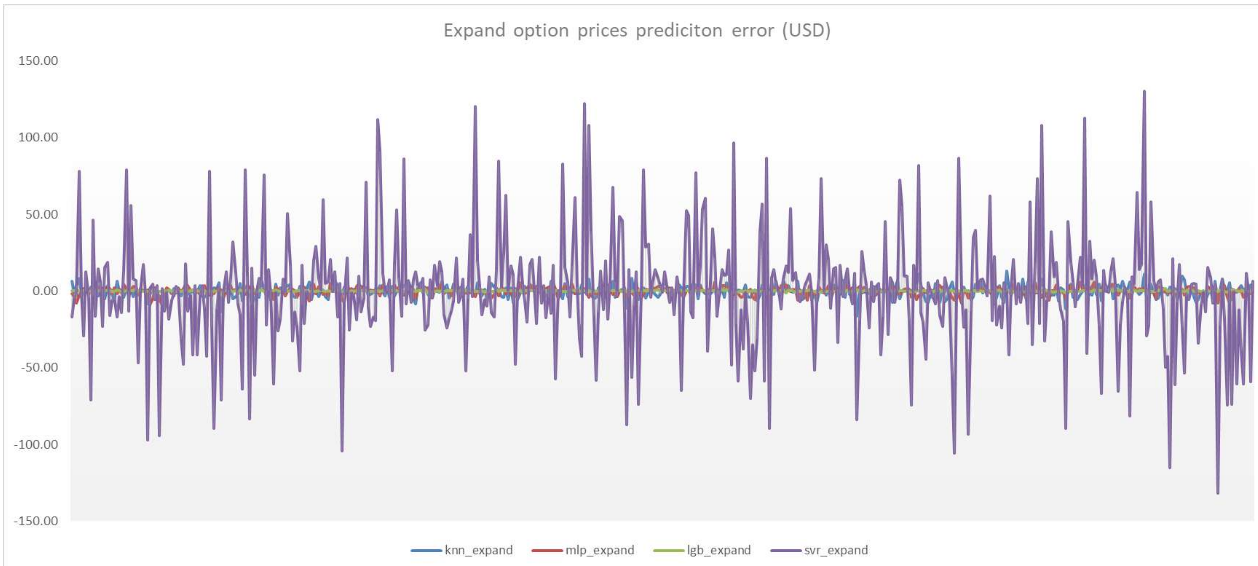
If investment project has a strategic value in it, at least one from above or some other advance real option is built into the project and its valuation becomes inevitable part of comprehensive analysis of value. Unfortunately, it’s widely known that Real Option valuation doesn’t account for the effect of competitors, and for those reasons is only valid for a monopolistic environment. Variety of techniques to account for competition within real option valuation framework has been developed and one example is to incorporate Game Theory discussed in [3]. While plausible, market competition may not follow equilibrium strategy derived by Game Theory models and then biases valuation even further. In this section, option to expand is computed for simulated investment projects data and machine learning models configured in previous section are trained to predict expansion option prices on unseen data. If concept of pricing real options with ML methods is proven, then training ML models on actual data is likely to solve monopolistic nature of real options pricing without need for explicit consideration of market competition effects.

Let's define formula for pricing real option using Cox-Ross-Rubinstein method of constructing binomial tree. Cone snippet for that is show next.

```
# function for option to expand
def option_to_expand(T,S,sig,r,N,expand,cost):

    dt=T/N
    dxu=math.exp(sig*math.sqrt(dt))
    dxd=1/dxu
    pu=((math.exp(r*dt))-dxd)/(dxu-dxd)
    pd=1-pu
    disc=math.exp(-r*dt)

    St = [0] * (N+1)
    V = [0] * (N+1)

    St[0]=S*dxd**N

    for j in range(1, N+1):
        St[j] = St[j-1] * dxu/dxd

    for j in range(1, N+1):
        V[j] = max(St[j]*expand-cost,St[j])

    for i in range(N, 0, -1):
        for j in range(0, i):
            V[j] = disc*(pu*V[j+1]+pd*V[j])

    return V[0]
```

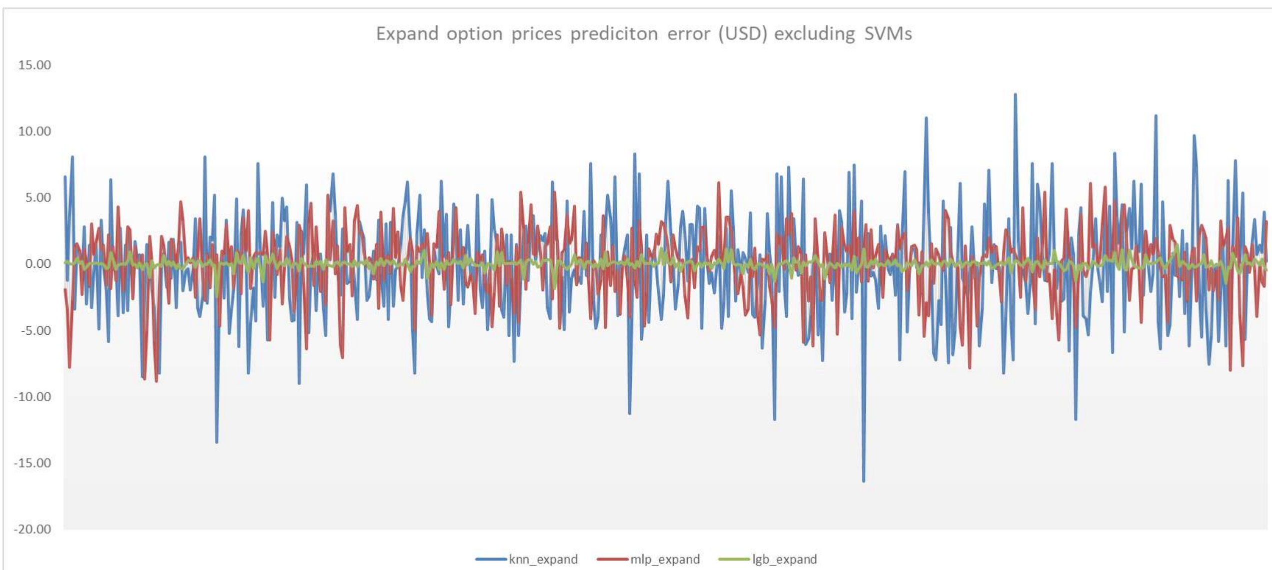
Next fixed interest rate of 5%, investment horizon of 3 years, and 1-year binomial steps are used to price option to expand with expansion factor of 1.5 at 50% cost of current project value. One thousand simulations of random numbers for projects' starting present value are drawn from range of values starting at 10 till 1000 with steps of 10, while volatility measures varying from 20 to 50% using 1% increments. As before, 5% of data is set aside for testing, and per each model prediction errors are graphed.

Figure 5 Prediction error for ML models on price of option to expand



Obviously, SVMs didn't lend themselves to accurately predict real option prices in the example but looking at the graph below excluding SVMs, it's clear that all other ML models seem to predict option prices very closely.

Figure 6 Prediction error for ML models on price of option to expand, excluding SVMs from the mix



Finally, let's look at average error and other prediction metrics in the table below.

Table 2 Options to expand pricing, prediction errors

metric	knn_expand	mlp_expand	lgb_expand	svr_expand
average error	3.09	1.93	0.26	24.03
min error	0.00	0.00	0.00	0.00
max error	16.36	8.80	2.45	131.65
max neg error	-16.36	-8.80	-2.45	-131.65
max pos error	12.81	6.12	1.48	130.13

Prediction error metrics suggest LGB as the best model for real options valuation.

Conclusion

This paper to some extent answered a question of whether Machine Learning models can be used to price financial options first and then real options. The former used real world data from historical prices Nasdaq Futures to compute option prices by The Black-Scholes model and then train ML models on them. The latter used simulated values of investment projects and their volatility estimates to compute predefined expansion option and then train ML models on them. Both experiments showed promising results with further room for improvements. If real option valuation using machine learning models can be statistically proven, strategic investment projects valuation formula could change from composite 1:

$$Project\ Value = Discounted\ Cash\ Flow + Real\ Options + Game\ Theory$$

To composite 2:

$$Project\ Value = Discounted\ Cash\ Flow + Machine\ Learning\ Option\ Price$$

Note that in composite 2, ML option price can already be trained to account for market competition.

References

[1] Robust Mean-Variance Hedging And Pricing of Contingent Claims In A One Period Model (2011) / *R. Tevzadze and T. UzunaShvili* / International Journal of Theoretical and Applied Finance / World Scientific Publishing Company

[2] Real Options Analysis: tools and techniques for valuing strategic investments and decisions / *Johnathan Mun* /2nd edition, published by John Wiley & Sons, Inc., 2006

[3] Investment Projects Valuation Using Real Options and Game Theory (2013) / *Levan Gachechiladze and Irakli Chelidze* / Academic-analytical Journal "Economics and Banking" 2013 Vol. I, N3

[4] Cox et al. (1979) / *Coxx, Ross, Rubinstein*

[5] Robustness (2008) / *Lars Peter Hansen, Thomas J. Sargent* / Published by Princeton University Press

[6] Hedging By Sequential Regression: An Introduction to the Mathematics of Option Trading (1989) / *H. FÄollmer and M. Schweizer* / ETH ZÄurich

[7] Financial Modeling, 3rd edition (2008) / *Simon Benninga* / MIT press

[8] Investment under Uncertainty (1994) / *Avinash K. Dixit and Robert S. Pindyck* / Princeton University Press

[9] Monte Carlo Methods in Financial Engineering/Paul Glasserman/published by Springer Science + Business Media, Inc., 2004

[10] Derivatives Market / Robert L. McDonald / published by Addison-Wesley, 2006

[11] Strategic Investment: real options and games / Han T.J. Smit and Lenos Trigeorgis / published by Princeton University Press, 2004

[12] Investment Projects Robust Valuation / Levan Gachechiladze, Revaz Tevzadze / MBA thesis, Georgian-American University, 2014

[13] Benninga, S., and M. Blume. 1985. On the Optimality of Portfolio Insurance. Journal of Finance, 40.

[14] Investment Projects Robust Valuation / Levan Gachechiladze, Tamaz Uzunashvili / Business Research Center of Georgian-American University, Applications of Random Processes and Mathematical Statistics in Financial Economics and Social Sciences II, Tbilisi Scientists and Innovations Festival 2017

[15] Application of Real Options in Valuation of Hedging Strategies / Levan Gachechiladze, Tamaz Uzunashvili, Teimuraz Toronjadze / Business Research Center of Georgian- American University, Applications of Random Processes and Mathematical Statistics in Financial Economics and Social Sciences III, Tbilisi Scientists and Innovations Festival 2018

[16] The Elements of Statistical Learning: Data Mining, Inference, and Prediction / Trevor Hastie, Robert Tibshirani, Jerome Friedman / Second Edition, Corrected 12th printing – Jan 13, 2017

[17] An introduction to Statistical Learning with Applications in R / Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani / Springer Science + Business Media New York, 2014

[18] Machine Learning for Algorithmic Trading, second edition / Stefan Jansen / Packt Publishing 2020

[19] Advances in Financial Machine Learning / Marcos Lopez de Prado / John Wiley and Sons, Inc. 2018

[20] Incorporating Machine Learning Models in Valuation of Hedging Strategies / Levan Gachechiladze, Tamaz Uzunashvili, Teimuraz Toronjadze / Business Research Center of Georgian- American University, Applications of Random Processes and Mathematical Statistics in Financial Economics and Social Sciences IV, Tbilisi Scientists and Innovations Festival 2021

[21] Application of Machine Learning in Option Pricing: A Review / Wends Li / Proceedings of the 2022 7th International Conference on Social Sciences and Economic Development (ICSSED 2022)

[22] Option pricing via machine learning / Tidy Finance with R, chapter 15 / Christoph Scheuch, Stefan Voigt, and Patrick Weiss / 2022.4-1

[23] Option pricing using Machine Learning / Codruț-Florin Ivașcu / Expert Systems with ApplicationsVolume 163, January 2021, 113799

[24] Option Pricing with Deep Learning / Alexander Ke, Andrew Yang / Stanford, SC230

[25] American option pricing with machine learning: An extension of the Longstaff-Schwartz method/ Jingying Lin, Caio Almeida / Brazilian Review of Finance, Vol. 19 No. 3 (2021)