دانشگاه صنعتی شریف

عنوان:

روش پارامتری عمیق معادلات دیفرانسیل جزئی و کاربردهای آن در قیمت‌گذاری اختیار خرید

اعضای گروه

احمد آقاپور
امیرمهدی حسین‌آبادی

نام درس

ریاضیات مالی

نیم‌سال دوم ۱۴۰۱-۱۴۰۲

نام استاد درس

دکتر مرتضی فتوحی
دکتر هیربد آسا

بسم الله الرحمن الرحیم

فهرست مطالب

فهرست تصاویر

چکیده: در این مقاله روش پارامتری عمیق معادلات دیفرانسیل جزئی را برای حل معادلات دیفرانسیل جزئی پارامتری در ابعاد بالا ارائه و بررسی میکنیم. تمرکز ما بر روی کاربردهای این روش در حوزه مالی می‌باشد. در این روش با استفاده از یک شبکه عصبی واحد، و بدون نیاز به داده‌ی نمونه، به حل خانواده‌ای از معادلات دیفرانسیل جزئی میپردازیم. بعد از یک دوره آموزش، جواب معادله به ازای هر مجموعه پارامتری در کمتر از میلی‌ثانیه قابل محاسبه است. به عنوان یک کاربرد واقعی، قیمت اختیارهای خرید و Greek هایشان را در یک مدل بلک_شولز چندمتغیره، به صورت بهینه محاسبه میکنیم. با آزمایش کردن روش در ابعاد مختلف، توانمندی روش پیشنهادی را نشان میدهیم.

واژه‌های کلیدی: معادلات دیفرانسیل جزئی پارامتری، قیمت گذاری اختیار خرید، شبکه عصبی عمیق، اختیار خرید سبدی

۱ مقدمه

هدف ما در این پروژه تکرار کردن روش ارائه شده در [۱] میباشد. بسیاری از مسائل در علوم مختلف با معادلات دیفرانسیل جزئی مدل‌سازی میشوند. حل معادلات دیفرانسیل جزئی پارامتری شده وقتی که ابعاد مسئله(تعداد پارامترها) افزایش می‌یابد بسیار دشوار می‌شود. اغلب روشهایی که برای این مسائل وجود دارند، دقت پایین دارند و یا کند هستند. در این روش برای حل معادلات دیفرانسیل جزئی از شبکه‌های عصبی عمیق[۱] استفاده میکنیم. مزیت شبکه‌های عصبی این است که با افزایش بعد محدود نمیشوند و کارایی خود را حفظ میکنند.

۱-۱ اهمیت موضوع

رویکرد ما برای حل این مسئله کاربردهای آن در دنیای مالی میباشد. خیلی از مسائل قیمت گذاری به فرم معادله دیفرانسیل جزئی سهموی مدل‌سازی میشوند. به طور خاص، حل مسئله قیمت گذاری اختیار خرید را مورد بررسی قرار میدهیم.

۱-۲ کلیات روش

در این مسئله، ابعاد مسئله(معادله دیفرانسیل) رابطه مستقیم با تعداد سهام زیرین اختیار خرید دارد. متغیرهای این مسئله زمان، حالت(قیمت سهام)، و پارامترهای دیگر مانند نرخ بازده بدون ریسک، نوسانات سهام، و همبستگی بین سهام می‌باشد.همه‌ی این متغیرها به عنوان ورودی مدل(شبکه عصبی عمیق) هستند و این مزیت بزرگ این روش محسوب میشود. چرا که پس از فاز یادگیری برون‌خط[۲]، در حالت برخط[۳]،میتوان جواب مسئله را به ازای هر زمان، حالت، و مجموعه‌ی پارامترها، در زمان بسیار کمی محاسبه نمود[۴]. شرط اولیه در این معادله میزان پرداخت در زمان سر رسید است که به قیمت توافقی[۵] وابسته می‌باشد.

روش کلاسیک برای آموزش مدل شبکه عصبی استفاده از یادگیری نظارتی[۶] می‌باشد. در این

[۱] Deep Neural Networks
[۲] Offline
[۳] Online
[۴] در اکثر مسائل دنیای واقعی که در حیطه مالی وجود دارند، مانند نظارت لحظه‌ای ریسک، بایستی محاسبات به صورت بهینه و در زمان بسیار کمی انجام شوند و مدل ارائه شده، به دلیل جامع و بهینه بودن، میتواند کاربرد زیادی در صنعت داشته باشد.
[۵] Strike Price
[۶] Supervised Learning

روش، برای تولید داده‌های آموزشی، بایستی جواب معادلات دیفرانسیل را برای تعداد زیادی از حالت‌ها محاسبه کنیم سپس با استفاده از بهینه سازی روی مقدار خطا، مدل را آموزش میدهیم. از آن جایی که در این روش نیاز به تعداد زیادی داده یادگیری[۷] داریم، پس نیاز به حل تعداد زیادی معادله دیفرانسیل با ابعاد بالا داریم که اصلا بهینه نمی‌باشد. به همین منظور، ما از یادگیری غیرنظارتی[۸] استفاده میکنیم. به این صورت که تخمینی از خطا بر حسب مقادیر خروجی مدل و گرادیان[۹] خروجی نسبت به پارامترها ارائه میدهیم و با اینکار، مدل بدون استفاده از داده نمونه، خودش آموزش می‌بیند[۱۰].

۱–۳ تازگی روش

کارهای جدید ما در این حوزه نسبت به سایر روش‌ها به این صورت میباشد:

۱. این روش خانواده‌ای از معادلات دیفرانسیل جزئی را با یکبار آموزش شبکه حل میکند. در این روش، معادله دیفرانسیل به طور کامل پارامتری میشود و همه‌ی پارامترها به عنوان ورودی به شبکه عصبی داده میشوند. با اینکار، بعد از یک فاز آموزش برون‌خط، در حالت برخط، جواب معادله به ازای هر دسته پارامتر ورودی در میلی‌ثانیه محاسبه میشود.

۲. با رویکردی کاربردی نسبت به مسئله، به محاسبه بهینه انواع مختلف اختیار خرید میپردازیم. برای اطمینان از مدل، به محاسبه خطای قیمت و نوسانان ضمنی در حالت‌های مختلف(برای مثال از ۱ تا ۸ سهم زیرین برای اختیار خرید را بررسی میکنیم) پرداختیم.

۳. به وسیله افزودن دانش پیشین در رابطه با قیمت گذاری اختیار خرید و تغییر مسئله معادله دیفرانسیل جزئی، دقت مدل را به طور محسوسی افزایش دادیم.

۱–۴ ساختار گزارش

ابتدا به معرفی برخی مفاهیم اولیه که در ادامه به آن‌ها نیاز پیدا میکنیم خواهیم پرداخت. سپس مدل ارائه شده را به صورت مفصل بررسی خواهیم کرد. سپس به نحوه‌ی پیاده‌سازی مدل گفته شده و تکرار خروجی‌ها که هدف اصلی پروژه بوده است میپردازیم و سپس نتایج حاصل شده را شرح میدهیم.

---

[۷] Training Data
[۸] Unsupervised Learning
[۹] Gradient
[۱۰] همگرایی این روش به جواب را نشان خواهیم داد.

۲ مفاهیم اولیه

در این فصل به معرفی مفاهیمی که در طول گزارش از آنها استفاده میشود میپردازیم.

۲-۱ اختیار خرید سبدی

اختیار خرید سبدی[۱۱]، مشابه اختیار خرید عادی، یک مشتق مالی[۱۲] است، با این تفاوت که دارایی زیرین آن، میانگین وزن‌دار سبدی از سهام میباشد. در زمان سر رسید، تابع پرداخت این اختیار خرید به این صورت میباشد:

$$G = \max\left\{ 0, \frac{1}{d}\sum_{i=1}^{d} s_i - K \right\} \tag{1}$$

که در اینجا s_i قیمت سهم iام در زمان سررسید است.

۲-۲ معادله دیفرانسیل جزئی سهموی

تابع $u(t,x)$ را در نظر بگیرید که t پارامتر زمان است و $x \in \Omega$ که $\Omega \in \mathbb{R}^d$ زیرمجموعه‌ای باز از فضای اعداد حقیقی است. در این صورت معادله‌ی زیر یک معادله‌ی دیفرانسیل جزئی سهموی[۱۳] میباشد:

$$\partial_t u(t,x) + L_x u(t,x) = f(t,x) \tag{2}$$

که در اینجا L یک عملگر دیفرانسیل جزئی به فرم زیر میباشد:

$$L_x u(t,x) = -\sum_{i,j=1}^{d} \partial_i(a_{ij}\partial_j u(t,x)) - \sum_{j=1}^{d} \partial_j(b_j u)(t,x) + cu(t,x) \tag{3}$$

در اینجا a_{ij}, b_i, c ضرایب هستند.

اگر پارامترهای موجود در معادله بالا را μ بنامیم میتوان نوشت:

$$\partial_t u(t,x;\mu) + L_x u(t,x;\mu) = f(t,x;\mu) \quad \mu \in P, \quad (t,x) \in Q = (0,T) \times \Omega \tag{4}$$

[۱۱] Basket Options
[۱۲] Financial Derivative
[۱۳] Parabolic Partial Differential Equation

که به آن معادله دیفرانسیل جزئی پارامتری [۱۴] گفته میشود.

در معادلات دیفرانسیل جواب‌ها معمولا یکتا نیستند و معمولا با شروط مرزی و اولیه مشخص میشوند. شرط مرزی نحوه‌ی رفتار تابع در مرز دامنه‌ای که در آن تعریف شده را مشخص میکند. شرط اولیه هم مشابه شرط مرزی است، با این تفاوت که برای نقاط مرزی زمان تعریف میشود. شرط اولیه به صورت:

$$\text{initial condition:} \quad u(\circ, x; \mu) = g(x; \mu) \;\; \forall x \in \Omega \tag{۵}$$

و شرط مرزی به صورت:

$$\text{boundary condition:} \quad u(t, x; \mu) = h(t, x; \mu) \;\; \forall (t, x) \in \Sigma = (\circ, T) \times \partial\Omega \tag{۶}$$

می‌باشد.

[۱۴] Parametric Parabolic PDE

۳ روش معادلات دیفرانسیل جزئی پارامتریک عمیق

۳-۱ کلیات روش

در قسمت مفاهیم اولیه، معادلات دیفرانسیل جزئی پارامتری را بررسی کردیم. میدانیم که بدست آوردن u به صورت صریح در خیلی از معادلات نشدنی میباشد. هدف ما در این روش، ارائه یک جواب تخمیمی برای u با استفاده از یک شبکه‌ی عصبی است. این شبکه عصبی مقادیر t, μ, x را به عنوان ورودی دریافت میکند و تخمینی از تابع $u(t, x; \mu)$ ارائه میدهد.

پارامتری کردن معادله دیفرانسیل این مزیت را دارد که دیگر نیازی به یادگیری شبکه‌های عصبی مختلف برای معادلات مختلف وجود ندارد. البته عیب این روش نیز این است که برای یادگیری یک مدل مناسب، به داده‌های خیلی بیشتری نیاز دارد و مدت زمان آموزش افزایش می‌یابد. در ادامه به بررسی قسمت‌های مختلف روش ارائه شده میپردازیم.

۳-۲ قیمت‌گذاری اختیار خرید چندمتغیره در مدل بلک-شولز

روش ارائه شده را میتوان برای تخمین جواب هر معادله دیفرانسیل جزئی به کار برد ولی در این پژوهش به معادلات مربوط به حوزه مالی یعنی قیمت‌گذاری اختیار خرید[۱۵] در مدل بلک-شولز[۱۶] پرداخته شده است. فرض میکنیم d دارایی زیرین[۱۷] در اختیار خرید وجود دارد. اگر قیمت سهام زیرین در فاصله زمانی t تا سررسید برابر با $s_i = e^{x_i}$ باشد، قرار میدهیم $x = (x_۱, x_۲, ..., x_d)$ و همچنین μ را برابر با پارامترهای دیگر یعنی نرخ بازده بدون ریسک[۱۸]،نوسانات سهام[۱۹] و همبستگی[۲۰] بین سهام قرار میدهیم. در این صورت $u(t, x; \mu)$ برابر با قیمت اختیار در لحظه $T - t$ میباشد. به عبارت دیگر میتوانیم بنویسیم:

$$u(t, x; \mu) = e^{-rt}\mathbb{E}(G(S_T(\mu))|S_t(\mu) = s) \tag{۷}$$

که در اینجا $S_t(\mu) = (S_t^۱(\mu), ..., S_t^d(\mu))$ قیمت سهام در لحظه t میباشد. قیمت سهام نیز از حرکت براونی[۲۱] مدلسازی میشوند. همچنین G تابع پرداخت[۲۲] در زمان سررسید میباشد. تابع پرداختی که

[۱۵] Option
[۱۶] Black-Scholes Model
[۱۷] Underyling Assets
[۱۸] Risk-free rate
[۱۹] Volatility
[۲۰] Correlation
[۲۱] Brownian Motion
[۲۲] Payoff Function

در این پژوهش از آن استفاده شده برابر است با:

$$g(x) = G(e^x) = (\frac{1}{d}\sum_{i=1}^{d} e^{x_i} - K)^+ = (\frac{1}{d}\sum_{i=1}^{d} s_i - K)^+ = max\left\{0, \frac{1}{d}\sum_{i=1}^{d} s_i - K\right\} \quad (8)$$

بدست آوردن مقدار u در مدل بلک_شولز معادل حل یک معادله دیفرانسیل جزئی میباشد. این معادله همگن[23] است یعنی $f(t,x;\mu) = 0$ و عملگر دیفرانسیل جزئی برابر است با:

$$L_x u(t,x;\mu) = ru(t,x;\mu) - \sum_{i=1}^{d}(r - \frac{\sigma_i^2}{2})\partial_{x_i} u(t,x;\mu) - \sum_{i,j=1}^{d} \frac{\rho_{ij}\sigma_i\sigma_j}{2}\partial_{x_i x_j} u(t,x;\mu), \quad (9)$$

دامنه‌ی متغیرها را یک ابرمکعب[24] در نظر میگیریم. در واقع برای هر متغیر یک بازه از اعداد حقیقی را به عنوان دامنه معرفی میکنیم.

۳-۳ تابع خطا

برای آموزش یک شبکه عصبی روی تعدادی داده‌ی نمونه، بایستی ابتدا یک تابع خطا معرفی کنیم و سپس پارامترهای شبکه عصبی را به گونه‌ای یاد بگیریم که تابع خطا کمینه شود. روش استفاده شده در اینجا، روش کمترین مربعات[25] میباشد. برای هر کدام از ۳ معادله موجود (معادله دیفرانسیل، شرط اولیه و شرط مرزی) یک تابع خطا تعریف میکنیم و با ترکیب این توابع یک تابع خطا کلی برای مدل ارائه میدهیم. مدل شبکه عصبی برای تخمین $u(t,x,\mu)$ را با $n(t,x,\mu)$ نشان میدهیم. در صورتی که میخواستیم از یادگیری نظارتی استفاده کنیم، می‌بایستی برای معادله دیفرانسیل اختلاف مقدار تابع u و n به ازای داده‌های نمونه به عنوان خطا در نظر میگرفتیم. ولی همانطور که قبلا هم توضیح دادیم، این روش مستلزم محاسبه مقدار $u(t,x;\mu)$ به ازای تعداد زیادی داده میباشد. به جای این روش ما از یادگیری غیرنظاراتی استفاده میکنیم و همانطور که در ادامه توضیح میدهیم، مقدار تابع خطا صرفا وابسته به تابع n میباشد. تابع خطای درونی را به این صورت تعریف میکنیم:

$$\mathcal{L}_{int}(n) = \frac{1}{|Q \times P|}\int_P \int_Q (\partial_t n(t,x;\mu) + L_x n(t,x;\mu) - f(t,x;\mu))^2 d_{(t,x)} d_\mu \quad (10)$$

و یا به عبارتی برابر میانگین جمع توان دوی خطاها میباشد. تابع خطا برای شرط اولیه برابر است با:

$$\mathcal{L}_{ic}(n) = \frac{1}{|\Sigma \times P|}\int_P \int_\Sigma (n(0,x;\mu) - g(x;\mu))^2 d_x d_\mu \quad (11)$$

[23] Homogeneous
[24] Hypercube
[25] Least Squares

برای سادگی در محاسبات از تابع خطای مربوط به شرط مرزی صرف نظر میکنیم. بدین صورت تابع خطای کل را برابر جمع توابع خطای بیان شده تعریف میکنیم:

$$\mathcal{L}(n) = \mathcal{L}_{int}(n) + \mathcal{L}_{ic}(n) \tag{۱۲}$$

برای آموزش شبکه عصبی، داده‌های تصادفی از توزیع‌های یکنواخت برای یادگیری غیرنظارتی تولید میکنیم:

$$(t^{(i)}, x^{(i)}, \mu^{(i)}) \in Q \times P, \quad (\hat{x}^{(i)}, \hat{\mu}^{(i)}) \in \Omega \times P, \quad i = ۱, ۲, ..., N \tag{۱۳}$$

سپس مقدار جملات خطا را به این صورت تعریف میکنیم:

$$\begin{aligned} \mathcal{L}_{int}(n) &= \frac{۱}{N}\sum_{i=۱}^{N}(\partial_t n(t^{(i)}, x^{(i)}; \mu^{(i)}) + L_x n(t^{(i)}, x^{(i)}; \mu^{(i)}) - f(t^{(i)}, x^{(i)}; \mu^{(i)}))^۲ \\ \mathcal{L}_{ic}(n) &= \frac{۱}{N}\sum_{i=۱}^{N}(n(۰, \hat{x}^{(i)}; \hat{\mu}^{(i)}) - g(\hat{x}^{(i)}; \hat{\mu}^{(i)}))^۲ \end{aligned} \tag{۱۴}$$

به این روشی که برای محاسبه‌ی انتگرال مقدار خطا از نمونه‌های تصادفی استفاده کردیم، روش مونت‌کارلو[۲۶] گفته میشود.

۳-۴ شبکه عصبی

معماری شبکه عصبی میتواند تاثیر زیادی در موفقیت مدل داشته باشد. مسائل مختلف نیاز به معماری‌های مختلف دارند. برای مثال شبکه‌ی عصبی پیچشی[۲۷] برای کاربردهای شناسایی تصویر استفاده میشوند و یا شبکه‌های حافظه‌ی بلند کوتاه مدت[۲۸] برای مدلسازی داده‌های ترتیبی کاربرد

[۲۶] Monte Carlo Integration
[۲۷] Convolution Neural Networks
[۲۸] Long Short-Term Memory(LSTM) Networks

دارند. معماری‌ای که در این روش استفاده شده است به این صورت است:

$$\begin{aligned}
S^{1} &= \sigma(W^{1}\overrightarrow{x} + b^{1}) \\
Z^{l} &= \sigma(U^{z,l}\overrightarrow{x} + W^{z,l}S^{l} + b^{z,l}), \quad l = 1, ..., L, \\
G^{l} &= \sigma(U^{g,l}\overrightarrow{x} + W^{g,l}S^{1} + b^{g,l}), \quad l = 1, ..., L, \\
R^{l} &= \sigma(U^{r,l}\overrightarrow{x} + W^{r,l}S^{l} + b^{r,l}), \quad l = 1, ..., L, \\
H^{l} &= \sigma(U^{h,l}\overrightarrow{x} + W^{h,l}(S^{l} \odot R^{l}) + b^{h,l}), \quad l = 1, ..., L, \\
S^{l+1} &= (1 - G^{l}) \odot H^{l} + Z^{l} \odot S^{l}, \quad l = 1, ..., L, \\
n(t, x; \mu) &= WS^{L+1} + b
\end{aligned} \tag{15}$$

به طوریکه که $\overrightarrow{x} = (t, x, \mu)$ ورودی‌های شبکه عصبی هستند. تعداد لایه‌های پنهان[29] برابر با $L + 1$ است و $\odot$ عملیات ضرب درایه‌ای[30] میباشد.تعداد واحدها در هر لایه برابر با M میباشد و $\sigma : \mathbb{R}^M \to \mathbb{R}^M$ یک تابع غیرخطی میباشد:

$$\sigma(x) = (\phi(z_1), ..., \phi(z_M)), \tag{16}$$

که در اینجا $\phi : \mathbb{R} \to \mathbb{R}$یک تابع فعالسازی غیر خطی[31] است. در این روش از tanh[32] استفاده شده است. پارامترهای شبکه عصبی را θ مینامیم که برابر است با نمام بردار ضرایب(وزن‌ها): $W^1 \in \mathbb{R}^{M\times(d+1)}$، $W \in \mathbb{R}^{1\times M}$، $(W^{*,l} \in \mathbb{R}^{M\times M})_{l=1}^{L}$ $(U^{*,l} \in \mathbb{R}^{M\times(d+1)})_{l=1}^{L}$، $b^1 \in \mathbb{R}^M$، $b \in \mathbb{R}$، $(b^{*,l} \in \mathbb{R}^M)_{l=1}^{L}$. آموزش شبکه عصبی معادل یافتن پارامترهای θ^* است به طوریکه:

$$\theta^* = argmin_\theta \mathcal{L}(n_\theta) \tag{17}$$

که به معنای یافتن پارامترهایی است که تابع خطا را کمینه کند. معماری ارائه شده در این مقاله برگرفته شده از [2] میباشد که یک معماری پیچیده میباشد چرا که درون هر لایه، تعدادی زیرلایه[33]وجود دارد. ویژگی مهم این معماری تکرار ضرب درایه‌ای بین توابع غیرخطی از ورودی میباشد. این به ما کمک میکند تا توابعی که به طور گسترده نسبت به ورودی‌ها(در اینجا زمان، قیمت سهام و سایر پارامترها) تغییر میکنند را مدلسازی کنیم. این معماری شباهت زیادی به معماری شبکه‌های بزرگراهی[34] و شبکه‌های حافظه‌ی بلند کوتاه مدت دارد.

---

[29] Hidden Layers
[30] Element-wise Multiplication
[31] Non-linear Activation Function
[32] Hyperbolic Tangent
[33] Sub-layer
[34] Highway Networks

۳-۵ استفاده از دانش پیشین

برای اینکه به دقت بیشتری در مدل برسیم، قیمت اختیار خرید را به ۲ قسمت تقسیم میکنیم: کران عدم آربیتراژ[۳۵]، و ارزش زمانی[۳۶]. کران عدم آربیتراژ برابر است با ماکسیمم صفر و امیدریاضی پرداخت اختیار خرید. به این صورت عمل میکنیم که شبکه عصبی به جای پیش‌بینی کل قیمت اختیار خرید، صرفا مقدار ارزش زمانی را پیش‌بینی میکند و با جمع کردن آن با کران عدم آربیتراژ که به صورت بسته قابل محاسبه است، به قیمت اختیار خرید میرسیم. پس قرار میدهیم:

$$u(t,x;\mu) = v(t.x;\mu) + \hat{u}(t,x;\mu) \tag{۱۸}$$

که در اینجا v همان ارزش زمانی اختیار خرید است و $\hat{u}$ نیز کران عدم آربیتراژ است و برابر است با:

$$\hat{u} = \max(\frac{\sum_{i=1}^{d} e^{x_i}}{d} - Ke^{-rt}, 0) \tag{۱۹}$$

با بیرون کشیدن این قسمت ثابت از معادله دیفرانسیل جزئی خواهیم داشت:

$$\begin{aligned} \partial_t v(t,x;\mu) + \mathcal{L}_x v(t,x;\mu) = \hat{f}(t,x;\mu), \quad \hat{f} = f - \partial_t \hat{u}, \quad \mu \in P, \quad (t,x) \in Q = (0,T)\times\Omega \\ v(0,x;\mu) = \hat{g}(x;\mu), \quad \hat{g} = g - \hat{u}, \quad x \in \Omega \\ v(t,x;\mu) = \hat{h}(t,x;\mu), \quad \hat{h} = h - \hat{u}, \quad (t,x) \in \Sigma = (0,T)\times\partial\Omega \end{aligned} \tag{۲۰}$$

مشکلی که $\hat{u}$، که در معادله دیفرانسیل هم ظاهر شده است دارد، این است که هموار[۳۷] نمیباشد. دلیل این ناهمواری وجود ماکسیمم در این تابع است. به همین منظور تخیمنی با استفاده از تابع سافت‌پلاس[۳۸] برای $\hat{u}$ ارائه میدهیم:

$$\hat{u}_\lambda(t,x;\mu) = \frac{1}{\lambda}\log(1 + e^{\lambda(\frac{1}{d}\sum_{i=1}^{d} e^{x_i} - Ke^{-rt})}), \quad 0 \leqslant \lambda \tag{۲۱}$$

این تابع تخمین خوبی برای $\hat{u}$ میباشد چرا که: $\lim_{\lambda\to\infty} \hat{u}_\lambda(t,x;\mu) = \max(\frac{\sum_{i=1}^{d} e^{x_i}}{d} - Ke^{-rt})$ و همچنین $\hat{u}_\lambda \in C^\infty$. در پیاده‌سازی قرار دادیم $\lambda = 0/1$ چرا که اگر مقدار آن را خیلی بزرگ قرار دهیم، مشتق دوم در زمان به روز رسانی شبکه عصبی خیلی بزرگ میشود و اگر هم مقدار آن را کوچک قرار دهیم، خطای تخمین خیلی زیاد خواهد شد.

---

[۳۵] no-arbitrage bound
[۳۶] time value
[۳۷] Smooth
[۳۸] Softplus Function

۳-۶ جزییات مربوط به قیمت‌گذاری اختیار خرید

۳-۶-۱ پارامترهای موجود در مدل

در این قسمت به بررسی پارامترهای مدل یعنی μ میپردازیم. میدانیم که:

$$\mu = \left\{r, K, (\rho_{ij})_{1\leqslant i,j\leqslant d}, (\sigma_i)_{i=1}^{d}\right\} \quad (22)$$ [۳۹]

که در بخش‌های قبلی آنها را بررسی کردیم.

میدانیم که قیمت‌سهام x و قیمت توافقی K واحد مشترک دارند(برای مثال دلار). همچنین با ضرب و تقسیم کردن همه‌ی این قیمت‌ها در یک عدد ثابت، تغییری در مسئله ایجاد نمیشود. به همین منظور، برای ساده‌سازی فرض میکنیم که در همه‌ی حالت‌ها داریم $K = 100$.

میدانیم که اگر تعداد سهام داخل سبد اختیار خرید برابر با d باشد، تعداد درایه‌های ماتریس همبستگی $(\rho_{ik})_{i,j}$ برابر d^2 میشود. این تعداد زیاد پارامتر پیچیدگی مدل را خیلی زیاد میکند . به همین منظور سعی میکنیم این ماتریس را به صورت بهینه پارامتری کنیم. میدانیم که ماتریس همبستگی بین سهام یک ماتریس متقارن[۴۰] و نیمه مثبت معین[۴۱] میباشد. یک رویکرد ساده میتواند استفاده از عوامل چولسکی[۴۲] باشد که منجر به پارامتری کردن با $\frac{d\times(d+1)}{2}$ پارامتر میشود که باز هم زیاد میباشد. رویکردی که ما استفاده میکنیم بر اساس روش کاربردی ارائه شده در [۳] میباشد. معمولا در مدل‌های قیمت‌گذاری، همه‌ی درایه‌های ماتریس همبستگی ظاهر میشوند ولی برای بدست آوردن ماتریس، اطلاعات محدودی در اختیار داریم. در این حالت، یک روش مناسب این است که بعضی از متغیرها را ثابت قرار دهیم و بقیه متغیرها را برون‌یابی[۴۳] کنیم. در روشی که در ادامه توضیح میدهیم، ماتریس را بر حسب تابعی از همبستگی متغیرهای(سهام) مجاور در ماتریس مشخص میکنیم:

$$\begin{aligned} \hat{\rho}_i &= \rho_{i,i+1} \in (-1,1) \\ \rho_{ji} = \rho_{ij} &= \prod_{k=i}^{j-1} \hat{\rho}_k \quad j > i \end{aligned} \quad (23)$$

و ثابت میشود که این ماتریس نیمه مثبت معین میباشد.

با نکاتی که گفته شد، یعنی ثابت کردن مقدار K و پارامتری کردن ماتریس همبستگی پارامترهای

---

[۳۹] در حالتی که اختیار خرید روی تک سهم باشد همبستگی ها حذف میشوند.
[۴۰] Positive
[۴۱] Positive Semi-definite
[۴۲] Cholesky Factors
[۴۳] Extrapolate

ورودی مسئله به این صوت میشوند:

$$\mu = \{r, (\sigma_i)_{i=۱}^{d}, (\hat{\rho}_i)_{i=۱}^{d-۱}\} \qquad (۲۴)$$

که در اینجا تعداد پارامترها برابر $۲d$ میباشد که با اضافه کردن tو x تعداد کل پارامترها $n = ۳d + ۱$ میشود.

۳-۷ قیمت‌گذاری‌های مرجع

در روش پیشنهادی، برای حالت‌های مختلف، قیمت‌گذاری اختیار خرید انجام میدهیم. در نتیجه، برای ارزیابی مدل، نیاز به قیمت‌گذاری‌هایی به عنوان مرجع داریم تا دقت مدل را بسنجیم. در ادامه برای انواع مختلف اختیار خرید، مدل‌های قیمت‌گذاری ارائه میدهیم.

۳-۷-۱ اختیار خرید اروپایی

این مسئله ساده‌ترین نوع قیمت‌گذاری اختیار خرید میباشد.طبق فرمول بلک‌ــ‌شولز میدانیم:

$$\begin{aligned} c(t, x; \mu) &= BS(t, e^x, r, \sigma, K) = \Phi(d_۱)e^x - \Phi(d_۲)Ke^{-rt} \\ d_۱ &= \frac{۱}{\sigma\sqrt{t}}(x - \ln(K) + rt + \frac{\sigma^۲ t}{۲}), \quad d_۲ = d_۱ - \sigma\sqrt{t} \end{aligned} \qquad (۲۵)$$

که یک جواب صریح برای اختیار خرید ارائه میدهد. ولی این فرمول برای اختیار خرید سبدی قابل استفاده نمی‌باشد.

۳-۷-۲ اختیار خرید سبدی با تابع پرداخت حسابی

در این حالت مسئله‌ی قیمت‌گذاری اختیار خرید سبدی (شامل d سهم زیرین) در یک مدل بلک‌ــ‌شولز چندمتغیره را داریم. از منظر عددی، قیمت‌گذاری چنین مشتقه‌ای نیازمند حل مسئله انتگرال‌گیری با ابعاد بالا از توابع ناهموار میباشد. به دلیل این ناهمواری، روش‌های انتگرال‌گیری عددی مرتبه بالا به طور مستقیم قابل استفاده نیستند. به همین دلیل از روش ارائه شده در [۴] استفاده میکنیم. در این روش با یک تبدیل متغیر، مسئله $d-$بعدی قیمت‌گذاری اختیار خرید را به ۲ مسئله‌ی ۱$-$بعدی و $(d-۱)-$بعدی هموار تبدیل میکنیم. در ادامه مراحل محاسبه قیمت اختیار خرید در این روش را به طور خلاصه بیان میکنیم:

۱. تجزیه ماتریس همبستگی مشابه روش گفته شده در [۴] و بدست آوردن مقادیر λ_i و $(v_{i,j})_{ij}$، به طوریکه برای $Y_i \in \mathcal{N}(۰, \lambda_i^۲)$، فرآیند تصادفی لگاریتم قیمت سهام به شکل زیر باشد:

$$\log(S_T^i(\mu)) = x_i + (r - \frac{\sigma_i^۲}{۲})t + Y_۱ + \sum_{j=۱}^{d} v_{i,j} Y_j, \tag{۲۶}$$

که در اینجا x_i لگاریتم قیمت سهم در فاصله زمانی t به زمان سررسید است.

۲. حل کردن یک امیدریاضی شرطی برای $Y_۱$ نسبت به $Y_۲, Y_۳, .., Y_d$، که منجر به تبدیل شدن مسئله قیمت‌گذاری اختیار خرید به یک مسئله $d - ۱$ بعدی با تابع پرداخت هموار میشود:

$$c(t, x; \mu) = \mathbb{E}(BS(۱, h(Y_۲, ..., Y_d), ۰, \lambda_۱, e^{-rt}K) \tag{۲۷}$$

که در اینجا $h(Y_۲, ..., Y_d) = \frac{۱}{d}\sum_{i=۱}^{d} e^{x_i - \frac{\sigma_i^۲ t}{۲}} e^{\sum_{j=۲}^{d} v_{i,j} Y_j}$. از آنجایی که تابع h هموار است و بعد مسئله یکی کاهش یافته، حل این مسئله نسبت به مسئله‌ی اولیه ساده‌تر است.

۳. برای حل مسئله بالا از روش GauB-Hermite که در [۵] ارائه شده است استفاده میکنیم.

۳-۷-۳ اختیار خرید سبدی با تابع پرداخت هندسی

برگرفته از [۲]، نوعی دیگر از اختیار خرید سبدی را نیز بررسی میکنیم. در این حالت، تابع پرداخت برابر است با:

$$G(x) = (e^{\sum_{i=۱}^{d} \frac{x_i}{d}} - K) \tag{۲۸}$$

به عبارت دیگر، میانگین هندسی سهام داخل سبد در نظر گرفته میشود. میانگین هندسی یک حرکت براونی هندسی چندمتغیره[۴۴] یک فرآیند تصادفی[۴۵] از همان جنس است و در نتیجه مسئله تبدیل به مسئله قیمت گذاری اختیار خرید معمولی با یک سهم که سود قابل پرداخت[۴۶] دارد میشود. برای سادگی در محاسبات، برای نوسانات سهام مقدار ثابت σ را در نظر میگیریم و برای همبستگی بین سهام نیز مقدار ثابت ρ را قرار میدهیم. سپس خواهیم داشت:

$$\begin{aligned} u(x, t; \mu) &= N(d_۱)e^{-qt}e^{\bar{x}} - N(d_۲)Ke^{-rt}, \\ d_۱ = \frac{۱}{\bar{\sigma}\sqrt{t}}(\bar{x} - \ln(K) + rt - qt + \frac{\bar{\sigma}^۲}{۲}t), &\quad d_۲ - d_۱ - \bar{\sigma}\sqrt{t}, \\ \bar{x} = \sum_{i=۱}^{d} \frac{x_i}{d}, \quad \bar{\sigma}^۲ &= \frac{\sigma^۲}{d(۱ + (d - ۱)\rho)}, \quad q = \sigma^۲ - \bar{\sigma}^۲ \end{aligned} \tag{۲۹}$$

[۴۴] Multivariate Geometric Brownian Motion
[۴۵] Stochastic Process
[۴۶] Dividend

۴-۷-۳ ارزیابی خطا با استفاده از نوسانات ضمنی

برای ارزیابی خطای مدل ارائه شده ۲ راه وجود دارد. در حالت اول خطای نسبی[۴۷] در قیمت را بررسی میکنیم. پس اگر پیش‌بینی مدل از قیمت اختیار خرید را u_{pred} بنامیم خواهیم داشت:

$$err_{rel}(t, s, \sigma, K) = \frac{|u_{pred}(t, s, \sigma, K) - u(t, s, \sigma, K)|)}{1 + |u(t, s, \sigma, K)|} \tag{۳۰}$$

حالت بعدی برای محاسبه خطا استفاده از نوسانات ضمنی[۴۸] است. در حالت ساده‌ی تک سهم، اگر قیمت دقیق سهم برابر با $u(t, x; \mu)$ باشد، نوسان ضمنی σ_{iv} را طوری تعریف میکنیم که:

$$BS(t, e^x, r, \sigma_{iv}, K) = u(t, x; \mu) \tag{۳۱}$$

که در اینجا BS تابع بلک_شولز می‌باشد. همچنین برای مدل پیشبینی $n(t, x; \mu)$ نیز مقدار نوسان ضمنی $\hat{\sigma}_{iv}$ را محاسبه میکنیم:

$$BS(t, e^x, r, \hat{\sigma}_{iv}, K) = n(t, x; \mu) \tag{۳۲}$$

سپس اختلاف مقدار σ_{iv} و $\hat{\sigma_{iv}}$ معیار سنجش خطا خواهد بود. برای تعمیم نوسان ضمنی به اختیار خرید سبدی، راه حل‌های مختلفی وجود دارد که در اینجا ما از روش ارائه شده در [۶] استفاده میکنیم که ساده‌ترین تعمیم میباشد. دراین روش مقدار σ_{iv} به این صورت بدست می‌آید:

$$BS(t, \sum_{i=1}^{d} \frac{e^{x_i}}{d}, r, \hat{\sigma}_{iv}, K) = n(t, x; \mu) = c \tag{۳۳}$$

که در اینجا c میتواند در بازه‌ی زیر قرار داشته باشد:

$$c_{lb} = (\sum_{i=1}^{d} \frac{e^{x_i}}{d} - Ke^{-rt}) \quad \leqslant \quad c \quad \leqslant \quad (\sum_{i=1}^{d} \frac{e^{x_i}}{d}) = c_{ub} \tag{۳۴}$$

دلیل اینکه نوسان ضمنی معیار مناسبی برای محاسبه خطا میباشد این است که مستقل از سبد سهام و کم و زیاد بودن قیمت آن‌ها، مقداری که برای نوسان ضمنی آنها بدست می‌آید بزرگی[۴۹] یکسانی دارند.

---

[۴۷] Relative Error
[۴۸] Implied Volatility
[۴۹] Magnitude

۴ پیاده‌سازی

برای پیاده سازی مدل ما از کتابخانه $Pytorch$ استفاده کردیم و آنرا در بستر $Colab$ ارائه شده توسط گوگل اجرا کردیم برای پردازش هر چه سریعتر از قابلیت موازی سازی و گرافیک استفاده کردیم این استفاده فقط مختص مدل یادگیری عمیق نبود بلکه برای حساب کردن قیمت دقیق اوراق اختیار خرید سبدی که از روش عددی استفاده میکرد نیز استفاده شد. این قسمت گفته شده با کمک کتابخانه $Cupy$ انجام شده که پیاده سازی کتابخانه معروف $Numpy$ به صورت موازی است .

۴-۱ پارامتر های مدل یادگیری عمیق

همانند آنچه در مقاله گفته شده ما از مدل سه زیرلایه ای استفاده کردیم که در هر زیرلایه اندازه هر لایه مخفی برابر ۹۰ میباشد همچین از بهینه ساز $Adams$ برای بهینه سازی شبکه استفاده کردیم و نرخ یادگیری برای این مدل را برابر ۰/۰۰۱ گذاشتیم. چنانکه از سخت افزار ضعیف تری نسبت به آنچه در مقاله استفاده شده است استفاده کردیم لذا مجبور بودیم مدل را در خطای بالاتری نسبت به خود مقاله متوقف کنیم ولی نتایج ما نتایج چندانی با نتایج مقاله نمیکرد

۴-۲ محاسبه مقدار دقیق اختیار خرید سبدی

ما همانند آنچه در مقاله قید شده بود از روش ارائه شده توسط [۴] استفاده کردیم ولی این روش با افزایش بعد نیازمند حافظه و زمان زیادی بود چرا که برای محاسبه انتگرال در یک بعد پایینتر از روش $GaussHermite\ quadrature$ استفاده میشد که نیازمند نمونه گیری زیاد در ابعاد بالا بود لذا ما برای فایق امدن بر این مشکل برای هر بعد فقط یکبار نمونه گیری کردیم و محاسبات را نیز توسط کتابخانه $cupy$ انجام دادیم تا هم حافظه کم نیاز داشته باشیم و هم مدت زمان کمتری برای این قسمت اختصاص داده شود.

۴-۳ محاسبه تلاطم ضمنی

با وجود روشهای مختلف برای محاسبه این قسمت و اشاره نشدن به روش خاص توسط مقاله ما از روش نیوتون–رافسون برای محاسبه استفاده کردیم

۵ نتایج

در این قسمت به بررسی حالت‌های مختلف قیمت‌گذاری اختیار خرید با ۱ تا ۸ سهم زیرین میپردازیم و خطای مدل در پیش‌بینی قیمت اختیار خرید و نوسان ضمنی را بررسی میکنیم. به صورت پیش‌فرض، بازه‌ی پارامترها به صورت:

$$\sigma_i \in [۰/۱, ۰/۳], \quad \hat{\rho}_i \in [۰/۲, ۰/۸], \quad r \in [۰/۰۱, ۰/۰۳], \quad T = ۴, \quad K = ۱۰۰, \quad s_i \in [۲۵, ۱۵۰] \tag{۳۵}$$

و مقادیر آن‌ها به صورت:

$$r = ۰/۰۲, \quad \sigma_i = ۰/۲, \quad \hat{\rho}_i = ۰/۵ \tag{۳۶}$$

خواهند بود.

برای همه‌ی مثال‌ها، از یک معماری واحد برای شبکه عصبی استفاده میکنیم که دارای ۹ لایه و در هر لایه ۹۰ راس است. در نتیجه برای استفاده از مدل در حالت‌های جدید نیازی به بدست آوردن ابرپارامترها نیست و این یک مزیت بزرگ این روش محسوب میشود.

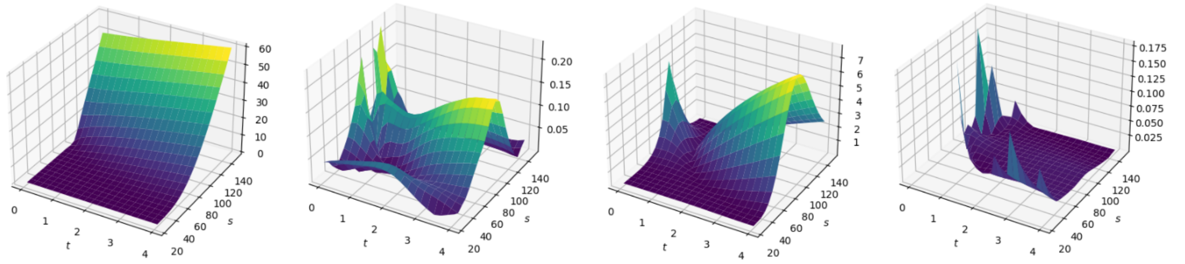
در ادامه به بررسی مسائل قیمت‌گذاری انواع مختلف اختیار خرید و مقادیر Greek آن‌ها میپردازیم.

۵-۱ اختیار خرید اروپایی با یک سهم زیرین

در ابتدا به بررسی اختیار خرید اروپایی با یک سهم زیرین میپردازیم و خروجی مدل را با جواب فرمول بلک-شولز مقایسه میکنیم. شکل ۱ نشان‌دهنده‌ی مقدار پیش‌بینی و خطای مدل برای پارامترهای ثابت می‌باشد. همانطور که در شکل مشخص است، خطای مطلق در پیش‌بینی قیمت در اکثر نواحی کمتر از ۰/۱ می‌باشد که نشان از دقت مدل است. در حالی که قیمت سهم در دامنه مورد بررسی بین ۰ و ۶۰ می‌باشد، مقدار تخمینی کران عدم آربیتراژ بخش زیادی از قیمت را به درستی توضیح میدهد. ارزش باقی‌مانده (ارزش زمانی) که توسط شبکه عصبی پیش‌بینی شده است، مقادیری در بازه‌ی $[-۱/۵, ۱۰]$ به خود گرفته است که در مقایسه با قیمت اختیار خرید کم می‌باشد. این نکته نشان از تاثیر جدی اعمال دانش پیشین در دقت مدل می‌باشد.

برای بررسی دقیق تر مدل، خطای نسبی نوسان نسبی را نیز مورد بررسی قرار میدهیم. میزان این خطا در اکثر نواحی کمتر از ۱٪ و در همه‌جا کمتر از ۲۰٪ می‌باشد. در نقاط گوشه‌ای نمودار که ناحیه‌هایی هستند که زمان باقی مانده تا سررسید کم است و یا فاصله‌ی قیمت سهم تا قیمت

توافقی زیاد است، مقدار تخمینی نوسان ضمنی دچار خطا میشود که در قسمت بعدی دلیل آن‌را شرح میدهیم.



شکل ۱: به ترتیب از چپ به راست: پیش‌بینی مدل از قیمت اختیار خرید، میزان خطای پیش‌بینی مدل از قیمت نسبت به قیمت واقعی، پیش‌بینی مدل برای ارزش زمانی اختیار خرید، و خطای نسبی نوسان ضمنی. در همه‌ی نمودارها، همه‌ی پارامترها (غیر از قیمت سهام) ثابت هستند.

۵-۲ اختیار خرید سبدی

در این قسمت به بررسی اختیار خرید سبدی با ۲ تا ۸ سهم میپردازیم. ابتدا به ارزیابی مدل در حالت‌هایی که پارامترها ثابت‌اند میپردازیم و سپس وابستگی مدل به پارامترها را بررسی میکنیم.

۵-۲-۱ ارزیابی برای پارامترهای ثابت

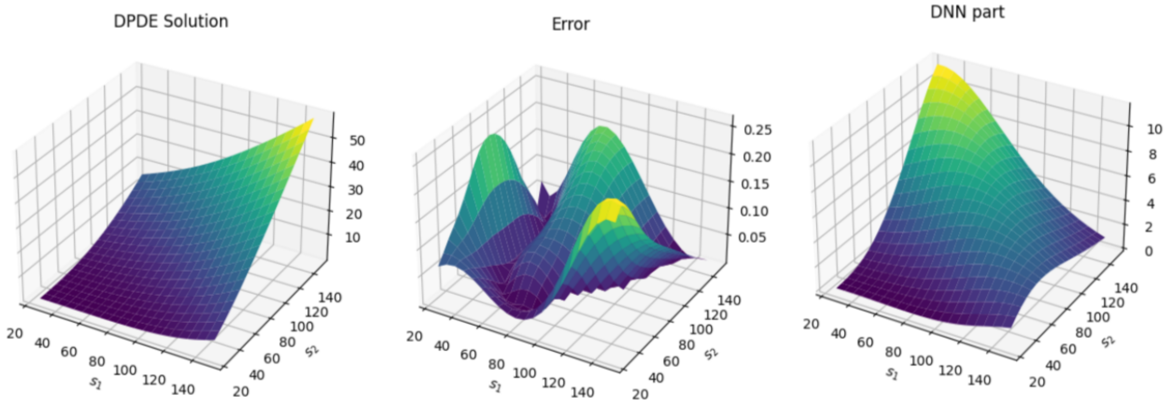
ابتدا سبدهای کوچک با ۲ و ۳ دارایی و سپس سبدهای بزرگتر با ۵ و ۸ دارایی را بررسی میکنیم. از آنجایی که نمایش توابع در ابعاد بالا دشوار است، در ابعاد مختلف نمودارهای متفاوت و متریک‌های مختلفی را مورد بررسی قرار میدهیم.

برای حالت دو بعد، پارامتر μ و $t = T$ را ثابت قرار میدهیم و قیمت سهام را متغیر قرار میدهیم. شکل ۲ نشان‌دهنده‌ی خروجی مدل و مقدار خطای آن در این حالت می‌باشد. همانطور که در شکل مشخص است، میزان خطا در تمام دامنه‌ی مورد بررسی زیر ۰/۲ میباشد. همچنین مشابه حالت تک‌سهم، مشاهده میکنیم که ارزش زمانی مقادیری در بازه‌ی [۰, ۱۲] به خود گرفته است که نشان از اهمیت ایده‌ی استفاده از دانش پیشین است.

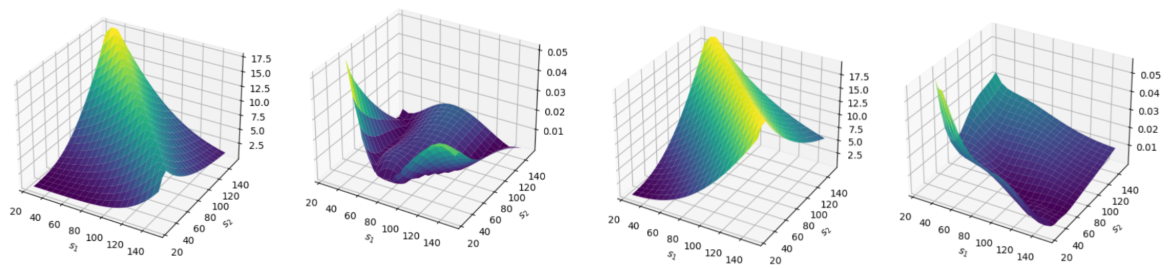
برای ارزیابی بهتر مدل، نوسانات ضمنی را نیز مورد بررسی قرار میدهیم. زمانی که مقدار $|s-K|$ یعنی فاصله قیمت دارایی با قیمت توافقی زیاد میشود، میزان تغییرات نوسان ضمنی نسبت تغییر

قیمت سهم حساسیت خیلی زیادی نشان میدهد. پس خطاهای گزارش شده در آن نواحی قیمتی نویزی خواهند بود. به همین دلیل، نوسان ضمنی را روی دامنه‌ی محدودتری که در ادامه توضیح میدهیم مورد بررسی قرار میدهیم و از حالت‌های مرزی صرف نظر میکنیم.

میدانیم که مقدار $u(t,x;\mu) - c_{lb}$ و $|s-K|$ با یکدیگر رابطه معکوس دارند. به زبان دیگر، در نواحی‌ای که اخلاف قیمت سهم با قیمت توافقی زیاد است، اخلاف قیمت اختیار خرید و کران عدم آربیتراژ افزایش می‌یابد. شکل ۳ گواه همین موضوع میباشد. در راستای صحبت قبلی در رابطه با محدود کردن دامنه بررسی نوسان ضمنی، مقدار خطای نسبی نوسان ضمنی حاصل از پیش‌بینی مدل را در فقط در نواحی‌ای که $u(t,x;\mu) - c_{lb} \geqslant ۰/۵$ نمایش میدهیم. دو شکل سمت راست و چپ نمودارهای مشابهی را برای دو پارامتر مختلف نشان میدهند. در هر دو حالتی که بررسی کردیم، مقدار خطای نسبی نوسان ضمنی در همه‌ی دامنه زیر ۵٪ میباشد و در اکثر نواحی نیز خیلی کمتر از این مقدار است که نشان‌دهنده‌ی دقت بالای مدل است.

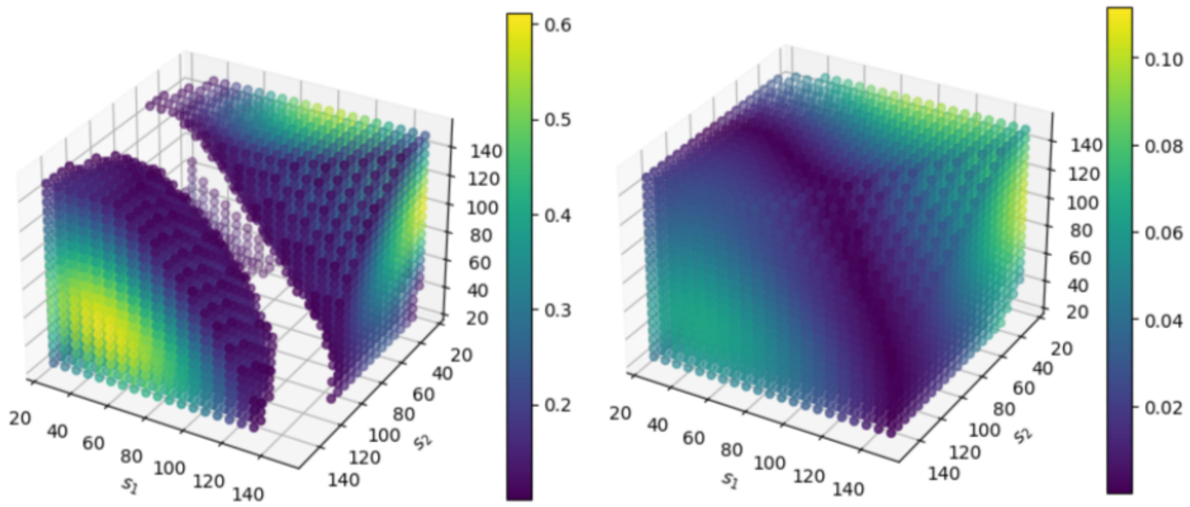


شکل ۲: به ترتیب از چپ به راست: پیش‌بینی مدل از قیمت اختیار خرید، میزان خطای پیش‌بینی مدل نسبت به قیمت واقعی، پیش‌بینی مدل برای ارزش زمانی اختیار خرید.



شکل ۳: دو شکل سمت چپ برای حالت $\sigma_۱ = ۰/۱, \sigma_۲ = ۰/۳$ میباشند و دو شکل سمت راست برای $\sigma_۱ = \sigma_۲ = ۰/۳$. شکل‌های اول و سوم از چپ نشان‌دهنده‌ی اختلاف قیمت اختیار خرید و کران عدم آربیتراژ $(u(t, x; \mu) - c_{lb})$ میباشند. همانطور که در شکل مشخص است، در گوشه‌های نمودار که اختلاف قیمت سهام با قیمت توافقی $K = ۱۰۰$ زیاد است، تابع رسم شده مقدار کمتری دارد. شکل‌های دوم و چهارم از چپ نشان‌دهنده‌ی میزان خطای نسبتی در حالتی که اختلاف قیمت اختیار خرید و c_{lb} بیشتر از ۰/۰۵ است.

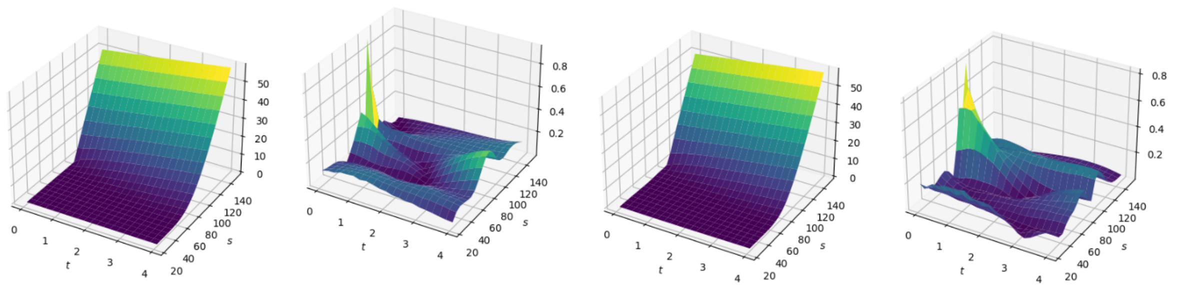
برای حالتی که ۳ سهم داریم نیز، پارامترهای t و μ را مشابه حالت ۲ سهم ثابت نگه میداریم و برای قیمت‌های مختلف، خروجی مدل را بررسی میکنیم. شکل ۴ نشان‌دهنده‌ی میزان خطا در قیمت و در نوسان ضمنی می‌باشند. در اکثر نواحی مقدار خطای قیمتی کمتر از ۰/۱ میباشد و حداکثر مقدار ۰/۶ به خود میگیرد. همچنین در تمام دامنه مورد بررسی، مقدار خطای نوسان ضمنی کمتر از ۰/۱، و در اکثر نواحی کمتر از ۰/۰۴ می‌باشد.



شکل ۴: خطاهای مختلف برای حالتی که ۳ سهم زیرین داریم. چپ: خطای موجود در قیمت اختیار خرید. فقط نواحی‌ای را نشان دادیم که میزان خطا بیشتر از ۰/۱ است. راست: خطای نسبی نوسان ضمنی در نواحی‌ای که اخلاف قیمت سهم با c_{lb} بیشتر از ۰/۵ است. پارامترها نیز برابر هستند با:

$\sigma_۱ = \sigma_۳ = ۰/۱, \sigma_۲ = ۰/۲, \rho_{۱,۲} = ۰/۲, \rho_{۲,۳} = ۰/۵$

برای بعدهای بالاتر از ۳، نمایش نمودارهای خطا مشابه حالت‌های قبلی غیرممکن می‌شود. به همین دلیل، همانطور که در شکل ۵ نشان داده شده است، میانگین پیش‌بینی قیمت اختیار خرید و حداکثر میزان خطا را روی مجموعه‌ای تصادفی از سبد سهام که میانگین برابر دارند، برای دو حالت $d = ۵, ۸$ بررسی میکنیم.. در نمودار دوم و چهارم از چپ مشاهده میکنیم که در اکثر نواحی مقدار خطا کمتر از ۰/۲ است.

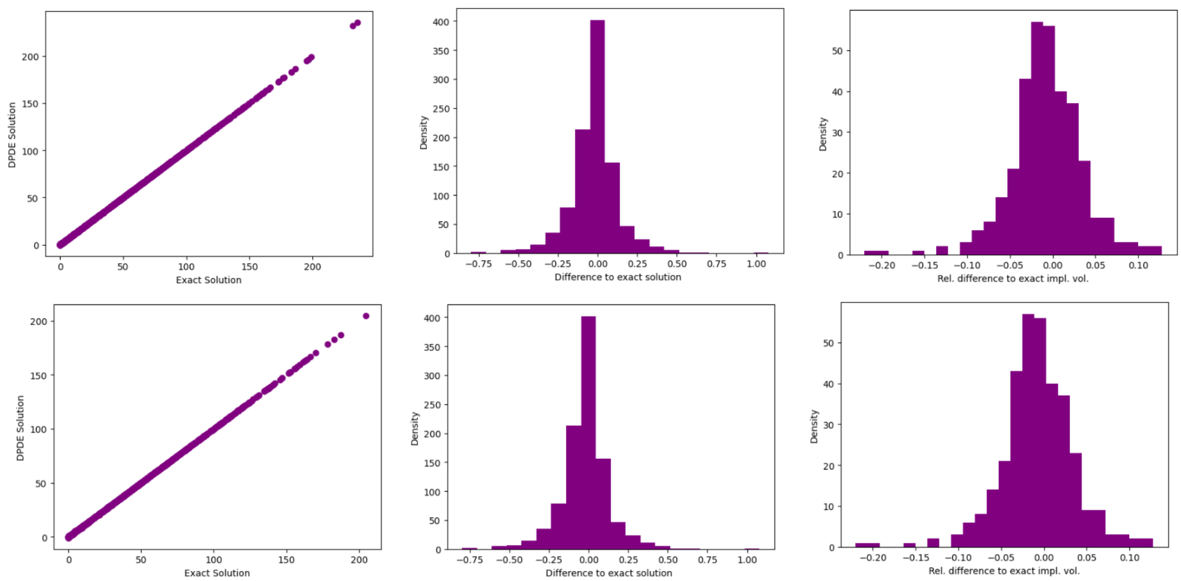


شکل ۵: نمودار نمایش قیمت پیشبینی و بیشینه خطا برای سبدهای سهامی که میانگین قیمت یکسانی دارند. دو شکل سمت چپ برای حالت ۵سهم و دو شکل سمت راست برای حالت ۷سهم می‌باشند.

۵-۲-۲ ارزیابی مدل در تمام دامنه پارامترها

در قسمت قبل، همه‌ی پارامترها غیر از قیمت سهام ثابت بودند. میدانیم که مزیت اصلی روش ارائه شده، حل مسئله به ازای تمامی پارامترها می‌باشد. در این قسمت به بررسی خطای پیش‌بینی برای پارامترهای متغیر میپردازیم. در حالتی که ۸ سهم در سبد داشته باشیم، مسئله ۲۵ بعدی می‌شود، و در نتیجه به نمایش کشیدن خطا دشوار میشود و باید از روش‌های دیگر استفاده کنیم.

در قدم اول، برای حالت $d = ۵,۸$ ، ۱۰۰۰ نقطه تصادفی از دامنه پارامترها تولید میکنیم و خروجی مدل را برای این مجموعه پارامترها محاسبه میکنیم. شکل ۶ نشان‌دهنده‌ی مقدار پیش‌بینی و دقیق جواب، خطای مطلق بین جواب دقیق و پیش‌بینی شده، و خطای نسبی نوسان ضمنی می‌باشند. در اکثر حالت‌ها مقدار خطای مطلق پیش‌بینی مدل از جواب دقیق کمتر از ۰٫۲ می‌باشد. هر چند داده‌های پرتی با خطای نزدیک ۰٫۸ نیز وجود دارند. برای نوسان ضمنی، تقریبا همه‌ی خطاها زیر ۱۰% میباشند. به طور کلی چیزی که مشخص است این است که عملکرد کلی مدل مناسب می‌باشد ولی از آنجایی که داده‌های پرت وجود دارند، برای بررسی بیشتر نحوه وابستگی مدل به پارامترها، آزمایش‌های دیگری نیز در ادامه انجام میدهیم.



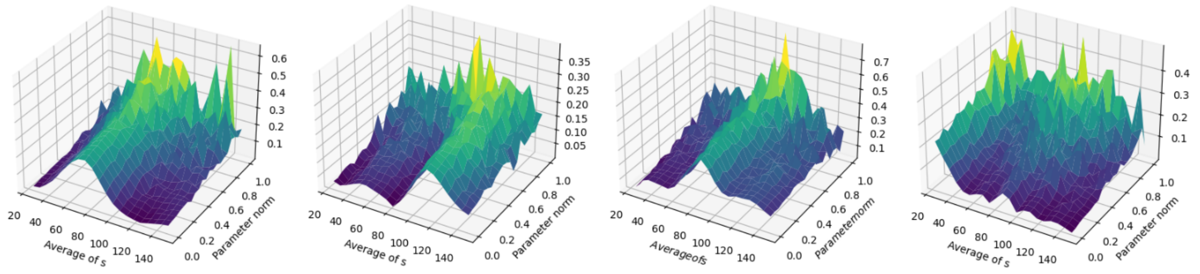
شکل ۶: نمودارهای بالا مربوط به حالت ۵ سهم و نمودارهای پایین مربوط به ۸ سهم می‌باشند. به ترتیب از چپ به راست: مقایسه‌ی مقدار پیش‌بینی و دقیق قیمت اختیار خرید، هیستوگرام مربوط به خطای مطلق قیمت پیش‌بینی، هیستوگرام خطای نسبی نوسان ضمنی. نمودارها برای ۱۰۰۰ نمونه تصادفی رسم شده‌اند.

در شکل ۷، بیشینه خطا برای نمونه‌هایی که میانگین قیمت سهام و نرم پارامتر برابر دارند را، در زمان $t = T$ و در ابعاد مختلف رسم کرده‌ایم. مقدار پارامترها نرمال‌سازی شده‌اند یعنی اینکه

$P = [-۱, ۱]^{n_\mu}$. در نتیجه مقدار نرم پارامترها نشان‌دهنده فاصله آن‌ها از پارامترهای مرجع یعنی $r = ۰/۰۲, \sigma_i = ۰/۲, \hat{\rho}_i = ۰/۵$ می‌باشد. همچنین نرم بی‌نهایت را در نظر میگیریم:

$$||\mu||_\infty = \max_{i \in \{۱,\dots,n_\mu\}} |\mu_i| \qquad (۳۷)$$

در شکل مشاهده میکنیم که با افزایش نرم پارامترها ، میزان خطا نیز افزایش می‌یابد، که نشان‌دهنده‌ی این است که در نواحی اطراف مرکز دامنه، مدل دقت بیشتری دارد.



شکل ۷: نمودار بیشینه خطا نسبت به نرم پارامترها و میانگین قیمت سهام در $t = T$. به ترتیب از چپ به راست برای حالت‌های $d = ۲, ۳, ۵, ۸$.

۵-۳ بررسی اختیار خرید سبدی با تابع پرداخت هندسی

همانطور که قبلا توضیح دادیم، در حالتی که تابع پرداخت اختیار خرید به فرم میانگین هندسی باشد، جواب صریح برای آن بدست می‌آید. در این حالت نحوه‌ی استفاده از دانش پیشین با حالت قبلی متفاوت است چرا که کران عدم آربیتراژ متفاوت است و برابر است با:

$$\hat{u}_\lambda(t, x; \mu) = \frac{۱}{\lambda} \log(۱ + e^{\lambda(e^{-Bt} e^{\sum_{i=۱}^{d} \frac{x_i}{d}} - Ke^{-rt})}) \qquad (۳۸)$$

حاصل تخمین مدل برای حالت ۲ سهم و پارامترهای ثابت و $t = T$ در شکل ۸ مشاهده میکنیم. همانطور که میبینیم، سطح خطاها کمتر از ۰/۱ است.

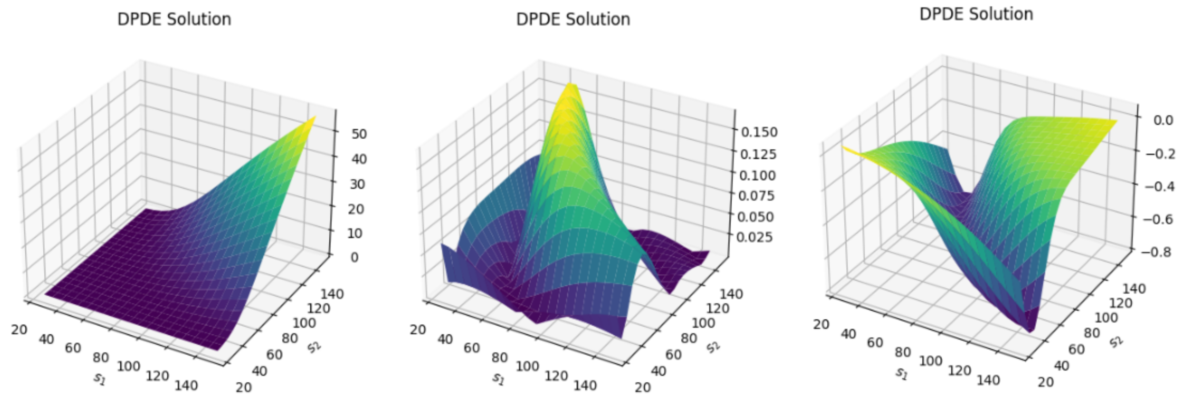
شکل ۹، بیشینه مقدار خطا را نسبت به نرم پارامترها و میانگین قیمت سهام در ۳ حالت $d = ۳, ۵, ۸$ را نشان میدهد. همانطور که در شکل مشخص است، اولا با افزایش نرم پارامترها خطا افزایش می‌یابد، دوما به طور کلی میزان خطا و بیشینه خطا با بعد مسئله رابطه مستقیم دارند.

برای محاسبه‌ی خطای نسبی، مشابه مسائل قبلی از نوسان نسبی استفاده میکنیم. در شکل ۱۰ هیستوگرام[۵۰] مربوط به اختلاف بین مقدار پیش‌بینی شده و دقیق نوسان نسبی را برای ۱۰۰۰ نمونه

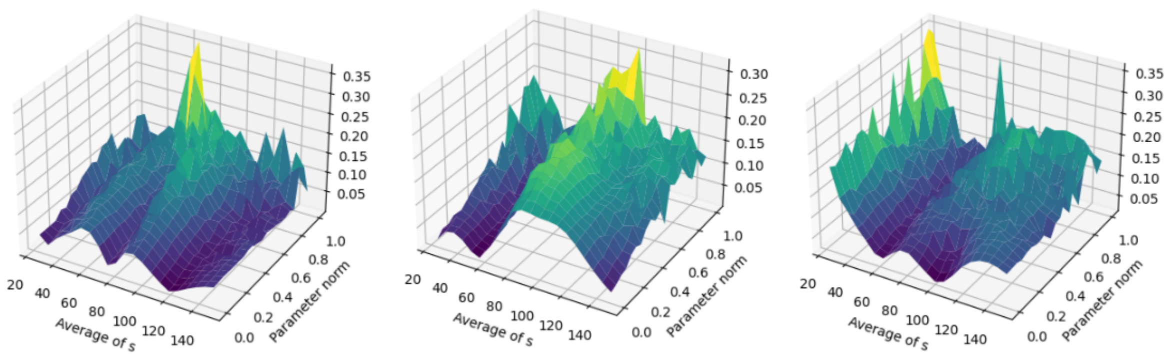
[۵۰] Histogram

تصادفی، برای هر کدام از حالت‌های $d = ۳,۵,۸$ نشان میدهد. مشاهده میکنیم که حتی در حالت $d = ۸$، خطای نسبی کمتر از ۰٫۰۵ است که نشان از دقت بالای مدل میدهد. به طور کلی میتوان گفت خطای نسبی نوسان ضمنی معیار بهتری برای سنجش مدل نسبت به خطای قیمتی می‌باشد.

به طور خلاصه، دیدیم که در این حالت از اختیار خرید نیز، مدل به خوبی پیش‌بینی میکند که تاییدی بر توانمندی[۵۱] مدل است.



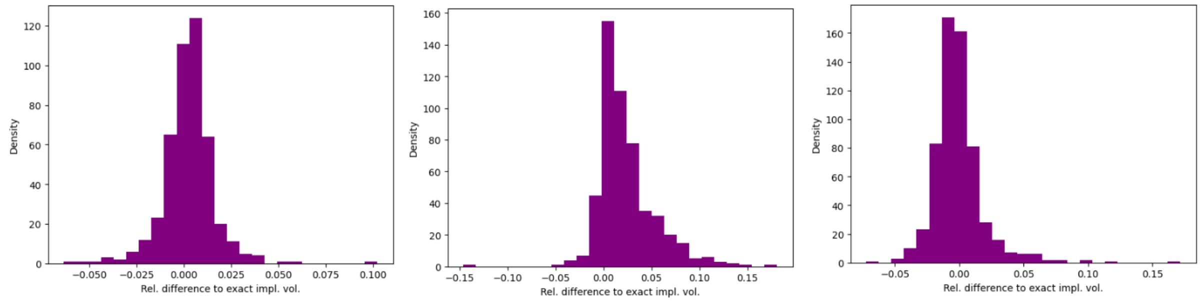
شکل ۸: به ترتیب از چپ به راست: پیش‌بینی مدل از قیمت اختیار خرید، میزان خطای پیش‌بینی مدل نسبت به قیمت واقعی، پیش‌بینی مدل برای ارزش زمانی اختیار خرید. همه‌ی عکس‌ها در حالت $t = T, \sigma = ۰٫۱$ میباشند.



شکل ۹: بیشینه خطا نسبت به نرم پارامترها و میانگین قیمت سهام در زمان $t = T$ در حالت تابع پرداخت هندسی. نمودارها از چپ به راست به ترتیب برای حالت های ۳، ۵، و ۸ سهم می‌باشد.

---

[۵۱] Robustness

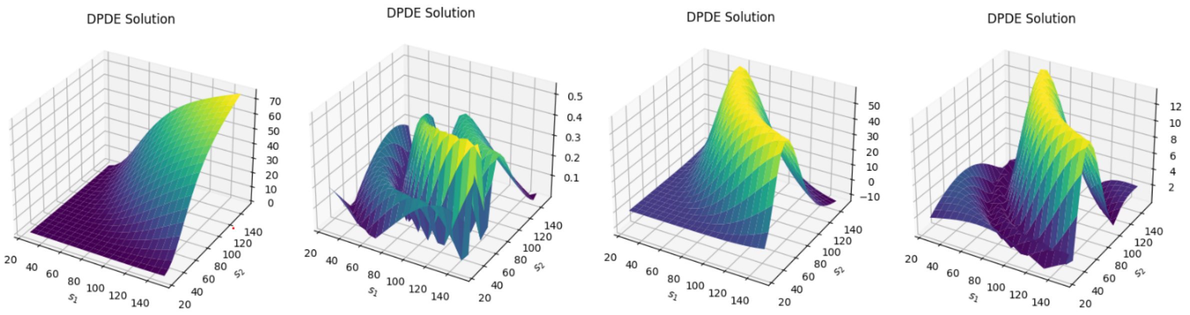


شکل ۱۰: هیستوگرام مربوط به اختلاف بین مقدار پیش‌بینی شده و واقعی نوسان نسبی برای ۱۰۰۰ نمونه تصادفی در حالت تابع پرداخت هندسی. نمودارها از چپ به راست به ترتیب برای حالت های ۳، ۵، و ۸ سهم می‌باشد.

۵-۴ Greeks

علاوه بر قیمت تخمینی، Greek ها نیز اهمیت زیادی در پوشش و مدیریت ریسک دارند. از آنجایی که در مدل ارائه شده، همه‌ی پارامترها به عنوان ورودی به شبکه عصبی داده شده‌اند، محاسبه‌ی مشتق خروجی نسبت به این پارامترها به راحتی قابل محاسبه است. شکل ۱۱، پیش‌بینی مدل از Δ[۵۲][۵۳] و Λ[۵۴] و خطای آن‌ها را در حالت $d = ۲$ نشان میدهد. خطای نسبی مدل در محاسبه Δ زیر ۲٪ است. از طرف دیگر خطای موجود در محاسبه Λ به طور قابل توجهی بیشتر می‌باشد. دلیل این مشاهده این است که مشتق قیمت اختیار خرید نسبت به قیمت سهام که همان Δ است به طور مستقیم در تابع خطای شبکه عصبی ظاهر میشود.

[۵۲] Delta
[۵۳] مشتق قیمت اختیار خرید نسبت به قیمت سهم اول زیرین.
[۵۴] Vega



شکل ۱۱: دو نمودار سمت چپ مربوط به پیش‌بینی و خطای مدل برای مشتق قیمت اختیار خرید نسبت به سهم شماره ۱ در بین سهام زیرین می‌باشد. دو نمودار سمت راست مربوط به پیش‌بینی و خطای مدل برای مشتق قیمت اختیار خرید نسبت به نوسانان قیمت سهام می‌باشد. در همه‌ی نمودارها $d = ۲$ می‌باشد.

References

[1] L. W. Kathrin Glau, "The deep parametric pde method and applications to option pricing," *Applied Mathematics and Computation*, 2022.

[2] K. S. Justin Sirignano, "Dgm: A deep learning algorithm for solving partial differential equations," *Computational Physics*, 2018.

[3] P. Dost, "Two useful techniques for financial modelling problems," *Applied Mathematical Finance*, 2010.

[4] M. S. CHRISTIAN BAYER and R. TEMPONE, "Smoothing the payoff for efficient computation of basket option prices," *Quant. Finance 18*, pp.491–505.

[5] P. C, "Function approximation for option pricing and risk management," *Queen Mary University of London*, 2020.

[6] A. G. . P. Tankov, "Implied volatility of basket options at extreme strikes," *Large Deviations and Asymptotic Methods in Finance*, 2015.

مطالب تکمیلی

در این قسمت توضیح مختصری از روش ارائه شده در [۴] ارائه میدهیم

لم ۱- *اگر ماتریس M یک ماتریس $n \times n$ مثبت معین باشد و v یک بردار دلخواه در فضای R^n باشد آنگاه ماتریس M را میتوان به صورت زیر تجزیه کرد*

$$M = VDV^T$$

که در آن ستون اول ماتریس V همان بردار v میباشد و D یک ماتریس قطری با درایه های نامنفی میباشد

اثبات. کافیست قرار دهیم $w = M^{-1}v$ و ماتریس زیر را بسازیم

$$\bar{M} = M - \frac{vv^T}{v^T w}$$

واضح است این ماتریس نیمه مثبت معین میباشد چرا که میدانیم برای هر دو بردار x, y داریم

$$||x^T y|| \leqslant ||A^{\frac{1}{2}} x|| ||A^{\frac{-1}{2}} y||$$

لذا کافیست ماتریس جدید ساخته شده را به صورت عمود تجزیه کنیم و واضح است که بردار v در فضای پوچ این ماتریس قرار دارد پس بقیه بردار ها عمود خواهند بود $\square$

با استفاده از لم بالا کافیست برادرv برابر بردار تماما یک قرار دهیم تا رابطه زیر را برای قیمت گذاری اختیار خرید سبدی داشته باشیم

$$\begin{aligned} C_{\mathcal{B}} &= E\left[\left(\sum_{i=1}^{d} w_i e^{(VY_i} - K\right)^+\right] \\ &= E\left[\left(\sum_{i=1}^{d} w_i \exp\left(Y_1 + \sum_{j=2}^{d} V_{i,j} Y_j\right) - K\right)^+\right] \\ &= E\left[\left(h\left(Y_2, \ldots, Y_d\right) e^{Y_1} - K\right)^+\right] \end{aligned}$$

که در آن Xلگاریتم طبیعی قیمت سهم در زمان پرداخت میباشد و V نیز ماتریس تلاطم میباشد لذا میتوانیم بنویسیم $X = VY$ و همچنین

$$h(\bar{y}) = h\left(y_2, \ldots, y_d\right) := \sum_{i=1}^{d} w_i \exp\left(\sum_{j=2}^{d} V_{i,j} y_j\right),$$

لذا میتوان نوشت

$$E\left[\left(\sum_{i=1}^{d} w_i e^{X_i} - K\right)^+ \mid \bar{Y}\right] = C_{BS}\left(h(\bar{Y})e^{\lambda_1^2/2}, K, \lambda_1\right),$$

و کافیست از این رابطه استفاده کنیم تا بعدمان را یک واحد کاهش دهیم

Abstract

We propose, formalise and analyse the deep parametric PDE method to solve highdimensional parametric partial differential equations with a focus on financial applications. A single neural network approximates the solution of a whole family of PDEs after being trained without the need of sample solutions. As a practical application, we compute option prices and Greeks in the multivariate Black–Scholes model as there is an urgent need for highly efficient methods. After a single training phase, the prices and sensitivities for different times, states and model parameters are available in milliseconds. By testing the method in different dimensions and option types, we confirm that the method is robust and efficient.

Keywords: Parametric Partial Differential Equations, Option Pricing, Deep Neural Networks, Basket Options,

Sharif University of Technology

Department of Mathematics

Mathematical Finance Project

The deep parametric PDE method and applications to option pricing

By:

Ahmad Aghapour

Amirmahdi Hosseinabadi

Course Instructors:

Dr. Morteza Fotouhi

Dr. Hirbod Assa

July 2023