بسم الله الرحمن الرحیم

عنوان مقاله:

Optimal Stopping Of Exotic American Option By Deep Learning

نفیسه ابراهیمی
محمدمعین شیرزاد

اساتید:
دکتر فتوحی، دکتر آسا، دکتر صلواتی و دکتر کاظمی

مرداد ۱۴۰۲

فهرست مطالب

۱ مقدمه

یک مسأله‌ی توقف بهینه (optimal stopping) شامل استفاده از متغیرهای تصادفی که به صورت متوالی مشاهده کرده‌ایم برای انتخاب یک زمان، به منظور اقدام برای انجام یک حرکت برای ماکسیمم کردن پاداش یا به صورت معادل، مینیمم کردن ضرر می باشد.
در این مقاله، این دنباله از متغیرهای تصادفی را با $X = (X_n)_{n=0}^{N}$ نشان می‌دهیم که یک فرآیند مارکف زمان گسسته که هر حالت آن، خود عضوی از $\mathbb{R}^d$ مقدار روی فضای احتمالاتی پالایش شده $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n=0}^{N}, \mathbb{P})$ است.
هدف ما توسعه دادن یک روش یادگیری عمیق می باشد به صورتی که بتواند به صورت مناسب و کارایی، سیاست بهینه (optimal policy) را، برای مسائل توقف به صورت زیر

$$\sup_{\tau \in T} \mathbb{E}[g(\tau, X_\tau)]$$

که در آن، $g : \{0, 1, \ldots, N\} \times \mathbb{R}^d \to \mathbb{R}$ توابع اندازه پذیر هستند و $\mathcal{T}$ نشان دهنده ی مجموعه ی تمام $X - stoppingtime$ها می باشد. برای این که اطمینان داشته باشیم که مسئله ی مذکور خوش تعریف است و دارای جواب بهینه می باشد، فرض می کنیم که تابع g انتگرال پذیر است و در شرط زیر (شرط اتگرال پذیری) صدق می کند.

$$\mathbb{E}[g(n, X_n)] < \infty \quad \text{برای همه ی} \quad n \in \{0, 1, \ldots, N-1\}$$

همچنین برای اینکه بتوانیم بازه ی اطمینان مقادیر بهینه ی مسئله را به دست آوریم، باید یک شرط قوی تری (شرط مربع-انتگرال پذیر (square-integrable)) نیز را فرض کنیم و آن شرط این است که:

$$\mathbb{E}[g(n, X_n)^2] < \infty \quad \text{برای همه ی} \quad n \in \{0, 1, \ldots, N-1\}$$

تعریف زمان توقف:
متغیر تصادفی را که مقادیر $\{0, 1, \ldots, N-1\}$ را به خود می گیرد، یک زمان توقف می گویم اگر برای هر $n \in \{0, 1, \ldots, N-1\}$، داشته باشیم $\{\tau = n\}$.

اگر بخواهیم در این چارچوب به صورت ریاضی بیان کنیم، می گوییم:
τ یک $X - stoppingtime$ است اگر $\{\tau = n\} \in \mathcal{F}_n$، $\forall n \in \{0, 1, \ldots, N\}$.
و مفهوم پاداشی که به دنبال ماکسیمم کردن آن هستیم، به صورت زیر می باشد:

$$V = \sup_{\tau \in T} \mathbb{E}[g(\tau, X_\tau)]$$

که در آن T مجموعه ی تمام $X - stoppingtime$ می باشد
و $g : \{0, 1, \ldots, N\} \times \mathbb{R}^d \to \mathbb{R}$ تابعی اندازه پذیر و انتگرال پذیر می باشد.

به صورت تئوری، مسائل توقف بهینه با تعدادی متناهی زمان توقف را می توان به صورت دقیق حل کرد، به صورتی که تابع ارزش آن توسط Snell envelope داده شده است و زمان توقف نظیر آن، اولین زمانی می‌باشد که پاداش توقف در حالت(state) فعلی، بیشتر از امید ریاضی پاداش ادامه دادن باشد.
در رویکرد عددی، این مسائل اغلب با برنامه ریزی پویا حل می شوند. اغلب رویکرد های مرسوم، در صورتی که بعد فرآیند مارکف بالا باشد، عملکرد بسیار ضعیفی دارند. این مشکلات، دامنه ی انواع مسائل توقف بهینه ی قابل حل را محدود می کنند. برای مثال، توانایی کار کردن در ابعاد بالا به ما این اجازه را می دهد تا قیمت اختیاری که دارای صدها سهام اساسی با زمان های سررسید متعدد می باشند، می دهد. همچنین، هر فرآیند ایکس فلانی را می توان با گنجاندن اطلاعات تاریخچه ی فرآیند در حالت یا وضعیت فعلی، به یک فرایند مارکف تبدیل کرد که البته به وضوح این کار باعث افزایش بعد را به همراه خواهد داشت.
یادگیری مایشن، رویکرید در حال رشد برای حل مسائل توقف بهینه می باشد. بسیاری از الگوریتم های یادگیری ماشین، بسیار قوی در حل مسائل در ابعاد بالا و همچنین به اندازه ی کافی منعطف، برای حل مسائل در حیطه های متنوع هستند. در این مقاله، تمرکز اصلی بر شبکه های عصبی چند لایه پیشخور (feed-forward multi-layer) می باشد که برای مسائل توقف بهینه به کار برده شده اند.
قبل از پرداختن به رویکرد اصلی، ابتدا به یادآوری تعدادی از مفاهیم می پردازیم.

۱.۱ تعاریف

قرارداد اختیار متعارف (option) :
دارنده ی یک قرارداد اختیار این اجازه ی این را دارد که یک مقدار مشخصی دارایی مالی را با در نظر داشتن یک زمان مشخص، با یک قیمت توافقی مشخص، بخرد یا بفروشد. هیچ لزومی برای اجرایی شدن قرارداد اختیار وجود ندارد. پس با توجه به تعریف قرارداد اختیار، در هر قرارداد اختیار چهار مورد باید مشخص شوند که عبارتند از:

۱. نوع قرارداد اختیار: قرارداد اختیاری که دارنده ی قرارداد اختیار، مقدار مشخصی دارایی را از صادرکننده ی قرارداد بخرد، قرارداد اختیار خرید (call option) نامیده می شود و قرارداد اختیاری که در آن دارنده ی قرارداد اختیار، مقدار مشخصی دارایی را به صادرکننده ی قرارداد بفروشد، قرارداد اختیار فروش (put option)] نامیده می شود.

۲. دارایی پایه (underlying asset) : که به صورت معمول می تواند سهام ریسکی (stock) ، سهام بدون ریسک (bond) و یا موارد مشابه دیگری باشند.

۳. مقدار دارایی پایه که معامله می شود.

۴. تاریخ انقضاء (expiration date) : اگر قرارداد اختیار را بتوان تا هر زمان قبل از زمان سررسید اجرایی کرد، این قرارداد اختیار، قرارداد اختیار آمریکایی (American option)

نامیده می شود ولی اگر قرارداد را فقط بتوان در زمان سررسید اجرا کرد، این قرارداد اختیار، قرارداد اختیار اروپایی (European option) نامیده می شود.

پس با توجه به تعاریف بالا، قرارداده های اختیار به صورت زیر تعریف می شوند:
قرارداد اختیار فروش اروپایی (European call option) :
دارنده ی قرارداد، حق دارد در زمان سررسید (T) ،یک واحد از دارایی پایه را به قیمت توافقی (stride price)، که معمولاً آن را با(k) نمایش می دهیم، به صادر کننده ی آن قرارداد بفروشد.

قرارداد اختیار خرید اروپایی (European call option) :
دارنده ی قرارداد، حق دارد در زمان سررسید (T) ،یک واحد از دارایی پایه را به قیمت توافقی (stride price)، که معمولاً آن را با(k) نمایش می دهیم، از صادر کننده ی آن قرارداد بخرد.

قرارداد اختیار فروش آمریکایی (American put option) :
قرارداد اختیار فروش آمریکایی دارنده ی قرارداد، حق دارد در هر زمان تا زمان سررسید (T)،یک واحد از دارایی پایه را به قیمت توافقی (stride price) ، که معمولاً آن را با(k) نمایش می دهیم، به صادر کننده ی آن قرارداد بفروشد.

قرارداد اختیار خرید آمریکایی (American call option) :
دارنده ی قرارداد، حق دارد در هر زمان تا زمان سررسید (T) ،یک واحد از دارایی پایه را به قیمت توافقی (stride price) ، که معمولاً آن را با(k) نمایش می دهیم، از صادر کننده ی آن قرارداد بخرد.

حال که با قرارداد های اختیار متعارف یا استاندارد آشنا شدیم، به آشنایی با قرارداد های اختیار نامتعلرف (exotic option) یا غیراستاندارد می پردازیم.

قرارداد اختیار نامتعارف (exotic option) :
قرارداد های اختیار نامتعارف، دسته ای از قرارداد های اختیار هستند که با قرارداد های اختیار مرسوم (قرارداد های استاندارد اروپایی و آمریکایی)، در شرایط فعال سازی، ساختار عایدی، تاریخ انقضاء و قیمت توافقی تفاوت دارند. به بیان دیگر، قرارداد های اختیار نامتعارف، همان قرارداد های اختیار اروپایی و آمریکایی هستند اما در آن ها شرط یا شرایط جدیدی لحاظ می شود که موجب رونق بیشتر این قرارداد های اختیار در خرید و فروش ها می شود. به سادگی، یک قرارداد اختیار نامتعارف، هر نوع قرارداد اختیاری غیر از قرارداد های اختیار متعارف در مبادلات عمده است. قرارداد های اختیار آسیایی و برمودا از این دسته قرارداد های اختیار هستند.

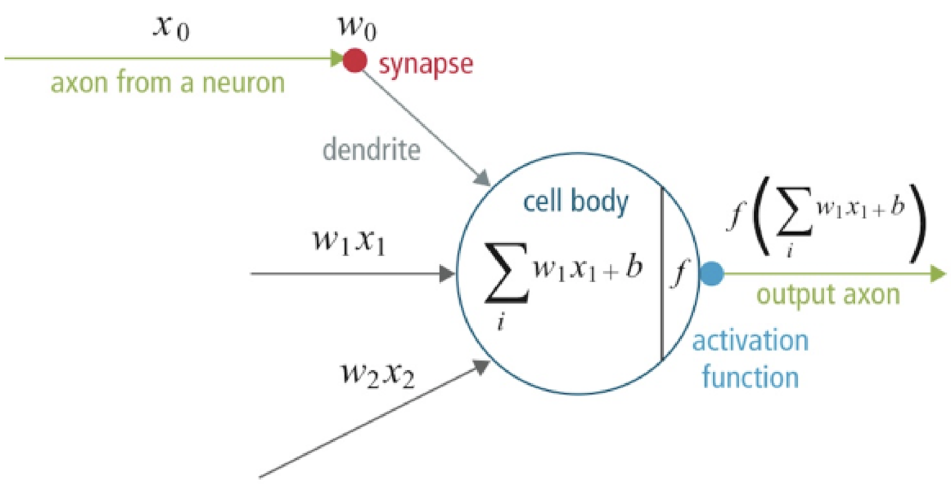
(Asian option)تعریف قرارداد اختیار آسیایی:
قرارداد اختیار آسیایی، اولین بار در سال ۱۹۸۷ در (Tokyo) معامله شد و به خصوص، در تجارت کالا ها بسیار محبوبیت دارد. در قرارداد اختیار آسیایی، (برخلاف قرارداد های اختیار اروپایی و آمریکایی که عایدی فقط وابسته به قیمت دارایی پایه در یک مقطع زمانی خاص می باشد)

عایدی وابسته به میانگین قیمت دارایی پایه در یک بازه ی زمانی خاص می باشد.

تعریف قرارداد اختیار برمودا (Bermudan) :
این نوع قرارداد اختیار، به نوعی، ترکیبی از قرارداد های اختیار اروپایی و آمریکایی می باشد. تاریخ اجرای این نوع قرارداد، در یک زیرمجموعه محدود از روز های مجاز که معاملات انجام می شوند و از قبل مشخص شده اند، می باشد.

۲.۱ شبکه های عصبی

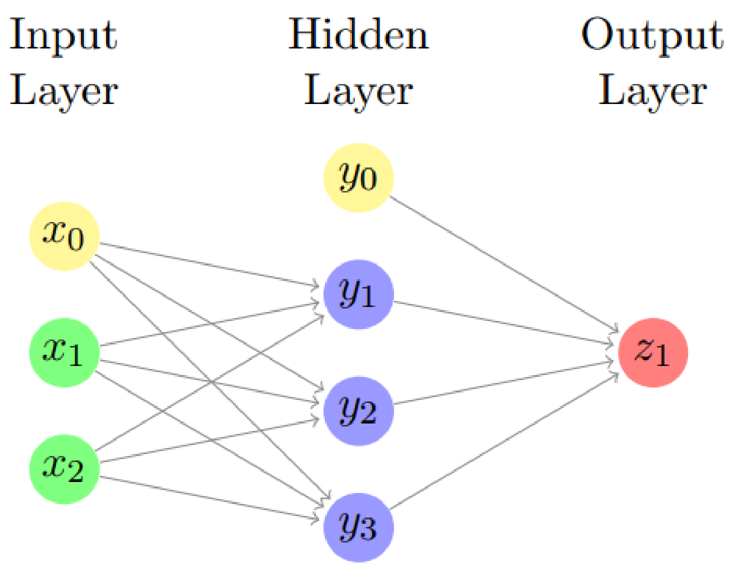
انگیزه ی اولیه برای استفاده از رویکرد ماشین لرنینگ در مسائل توقف بهینه، توانایی آن ها برای (curse of dimensionality) غلبه بر نفرین بعد می باشد که روش های عددی مرسوم با آن مواجه بودند. شبکه ی عصبی یک سیستم محاسباتی می باشد که به صورت ساده سازی شده بر اساس مغز انسان طراحی شده است. بدون عمیق شن در جزئیات زیستی، ایده اصلی این است که یک نرون ورودی را از نرون های دیگر یا از یک منبع خارجی دریافت می کند و از آن برای تولید خروجی استفاده می کند. این خروجی، می تواند به عنوان ورودی به نرون بعدی داده شود و یا به عنوان خروجی کل سیستم در نظر گرفته شود. در شبکه عصبی محاسباتی، به این نرون ها، نودهای شبکه عصبی گفته می شود.
روشی که در آن یک نود ورودی را دریافت می کند و آن را به خروجی تبدیل می کند، در شکل (۱) نشان داده شده است.



شکل ۱: یک نود در شبکه عصبی

x_1 و x_2 ورودی ها هستند، که اگر این نود در لایه اول باشد، این ها ورودی کل شبکه می باشند و اگر این نود در لایه اول نباشد، این x_1 و x_2 خروجی های لایه ی قبلی می باشند. هر

x_i دارای یک w_i متناظر می باشد که وزن آن نامیده می شود و اهمیّت نسبی ورودی مورد نظر را کمّی می کند. سپس، همان طور که در شکل(۱) دیده می شود، مجموع ورودی ها و وزن ها با یک ترم اضافه بایاس جمع می شوند. این مجموع، توسط تابعی به نام تابع فعال ساز (activation function) که با f نمایش داده شده است، دچار تغییراتی می شود و پس از آن، تبدیل به خروجی نود می شود. خروجی تابع فعال ساز را می توان به عنوان (firing rate) نسبی نود در نظر گرفت.
شکل (۲) چگونگی اتصال نود ها در معماری یک شبکه عصبی متراکم (dense) ، تک لایه ای (single-layer) پیشخور (feed-forward) را نشان می دهند. اصطلاح متراکم به معنای این است که نود ها به صورت کامل به یگدیگر متصل هستند؛ یعنی یک نود از تمام نود های لایه ی قبلی، ورودی دریافت می کند و خروجی خود را به تمام نود های لایه ی بعدی می فرستد. تک لایه ای بودن نیز به این معنی می باشد که تنها یه لایه نهان در شبکه عصبی وجود دارد.



شکل ۲: مثالی از یک شبکه عصبی پیشخور

برای درک بهتر آنچه در شبکه عصبی شکل به صورت ریاضی انجام می شود.

 مورد ۴

۲ صورت مسأله

در این بخش، روی چارچوب ریاضیاتی مورد نیاز برای پیاده سازی یک شبکه عصبی چند لایه (multi-layer) پیشخور (feed-forward) برای حل مسائل توقف بهینه به فرم زیر تمرکز می کنیم.
X دنباله ای از متغیرهای تصادفی می باشد که آن را با $X = (X_n)_{n=0}^{N}$ نمایش می دهیم و یک فرآیند مارکف زمان گسسته (و با افق متناهی) دی حقیقی مقدار روی فضای پالایش شده احتمالاتی $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n=0}^{N}, \mathbb{P})$ است.
می گوییم τ یک $X - stoppingtime$ است اگر $\{\tau = n\} \in \mathcal{F}_n$، $\forall n \in \{0, 1, \ldots, N\}$.
و مفهوم پاداشی که به دنبال ماکسیمم کردن آن هستیم، به صورت زیر می باشد:

$$V = \sup_{\tau \in T} \mathbb{E}[g(\tau, X_\tau)] \tag{1}$$

که در آن T مجموعه ی تمام $X - stoppingtime$ می باشد
و $g : \{0, 1, \ldots, N\} \times \mathbb{R}^d \to \mathbb{R}$ تابعی اندازه پذیر و انتگرال پذیر می باشد.

۱.۲ صورت بندی ریاضی مسأله توقف بهینه

صورت بندی ریاضی مسأله توقف بهینه دارای ۴ بخش می باشد:

۱. فرآیند تصادفی:
فرآیند تصادفی مارکف X که زمان گسسته با افق متناهی می باشد به صورتی که

$$X = (X_n)_{n=0}^{N}\,, \quad X_n \in \mathbb{R}^d\,, \quad \forall n \in \{0, 1, \ldots, N\}$$

۲. تابع پاداش یا عایدی (pay off) :

$$g(t, X_1, \ldots, X_N) \xrightarrow{\text{با توجه به مارکف بودن فرآیند}} g(\tau, X_\tau)$$

۳. هدف (objective) :

$$\sup_{\tau \in T} \mathbb{E}[g(\tau, X_\tau)]$$

۴. استراتژی های مجاز:
در استراتژی های مجاز، نباید از اطلاعات آینده برای تصمیم گیری استفاده کرد.

حال که به صورت کامل با ساختار مسأله آشنا شدیم، بر روی چارچوب ریاضیاتی مورد نیاز برای پیاده سازی شبکه عصبی چند لایه (multi-layer) پیشخور (feed-forward) برای حل مسائل توقف بهینه تمرکز می کنیم. برای این کار، ابتدا باید به اثبات دو قضیه ی کلیدی بپردازیم. قضیه ی اول نشان می دهد که زمان توقف بهینه، τ، را می توان به صورت تابعی از یک سری تصمیمات توقف صفر و یکی بیان کرد. قضیه ی دوم نشان می دهد که می توانیم این تصمیمات توقف صفر و یکی را می توان با یک شبکه عصبی چند لایه پیشخور تقریب زد. این نتایج کلیدی مبنای ریاضی برای اینکه چگونه می توان مسائل توقف بهینه را با شبکه عصبی حل کرد، می باشند.

۲.۲ بیان کردن زمان های توقف به صورت یک سری از تصمیمات توقف

اولین کاری که باید انجام بدهیم این است که نشان دهیم تصمیم به توقف فرآیند مارکف را $(X_n)_{n=0}^N$ می توان بر اساس دنباله ای از توابع $\{f_n(X_n)\}_{n=0}^N, f_n : \mathbb{R}^d \to \{0,1\}$ گرفت. برا ی انجام این کار، در ابتدا، مسأله ی کلی توقف بهینه ی خود در ۱ را به دنباله ای از مسائل توقف تقسیم می کنیم. به طور خاص، برای زمان n، مسئله ی توقف را به صورت زیر در نظر بگیرید:

$$V_n = \sup_{\tau \in \mathcal{T}_n} \mathbb{E}[g(\tau, X_\tau)] \tag{۲}$$

که در آن، $\mathcal{T}_n$ مجموعه ی تمام $X - stoppingtime$ ها به صورتی که $n \leq \tau \leq N$ می باشد، است. به وضوح، از آن جایی که فرآیند باید در زمان N توقف کند، $\mathcal{T}_N = \{N\}$ و بنابراین زمان N، زمان توقف بهینه می باشد و $\tau_N = N$. علاوه بر این، ازآن جایی که $\{f_n(X_n)\}_{n=0}^N$ دنباله ی صفر و یکی تصمیمات توقف ما می باشد، باید ذکر شود که $f_N \equiv 1$ و لذا می توان به این صورت $\tau_N = N f_N(X_N)$ نوشت.
حال که زمان توقف زمان N را به صورت تابعی از تصمیمات توقف صفر و یکی نوشتیم، قصد داریم این کار را برایn هایی که $0 \leq n \leq N-1$ نیز انجام دهیم. به این معنی که قصد داریم زمان توقف n، τ_n، را به صورت تابعی از $\{f_k(X_k)\}_{k=n}^N$ بنویسیم. قصد داریم با استفاده از معادله زیر انجام دهیم:

$$\tau_n = \sum_{k=n}^{N} k f_k(X_k) \prod_{j=n}^{k-1} (1 - f_j(X_j)) \tag{۳}$$

که در آن $f_N \equiv 1$ می باشد. معادله فوق یک زمان توقف درτ_n را تعیین می کند. البته این که این زمان توقف، یک زمان توقف بهینه می باشد را هنوز تأیید نکرده ایم. قضیه زیر نشان می دهد که τ_n یک زمان توقف بهینه در زمان n می باشد.

۳.۲ قضیه (۱)

در این مقاله، به طور کلی، دو قضیه ی مهم را ثابت می کنیم. قضیه ی اول، بهینه بودن انتخابمان را نشان می دهد و قضیه ی دوم، کارایی و اثربخشی مدل را اثبات می کند.
در اثبات قضیه ی یک از یک لم استفاده می شود که در ابتدا به بیان و اثبات آن لم می پردازیم و سپس به سراغ اثبات قضیه ی یک می رویم. نام لم پایین، لم Doob-Dynkin می باشد که دارای نسخه های متعدد متفاوتی می باشد. ایده ی اصلی لم، به طور کلی، لم Doob-Dynkin ، شرایطی روی دو تابع X و Y در نظر می گیرد به صورتی که این شرایط، وجود یک تابع مانند ϕ به صورتی که $X = \phi \circ Y$ است را تضمین می کند.

$$\begin{tikzcd} & \Omega_X \\ \Omega \arrow[ur, "X"] \arrow[dr, "Y"'] & \\ & \Omega_Y \arrow[uu, "\phi"'] \end{tikzcd}$$

شکل ۳: شمای کلی لم Doob-Dynkin

آن نسخه از لم که در اثبات استفاده شده است، Conditional expectation Doob-Dynkin می باشد و به شرح زیر است.

لم ۱ (Conditional expectation Doob-Dynkin). فرض کنید $\Gamma : \Omega \to [0, \infty]$ و $Y : \Omega \to \Omega_Y$ اندازه پذیر باشند. اگر Y یک $\sigma - finite$ با $\sigma - field(\sigma - algebra)$ اولیه $\mathcal{E}_Y$ باشد، آن گاه یک تابع اندازه پذیر یکتا مانند $\phi : \Omega_Y \to \Omega_\Gamma$ وجود دارد به صورتی که $\mathbb{E}(\Gamma|\mathcal{E}_Y) = \phi(Y)$.

نتیجه ی حاصل از قضیه ی زیر نشان می دهد که در روش ما که در آن از روش بازگشتی به منظور محاسبه ی یک تقریب از جواب مسئله ی توقف بهینه استفاده می کنیم، کافی است که زمان توقف مسئله را به فرم (۳) در نظر بگیریم.

قضیه ۱. برای یک $n \in \{0, \dots, N-1\}$، فرض کنید τ_{n+1} یک زمان توقف در $\mathcal{T}_{n+1}$ ($\tau_{n+1} \in \mathcal{T}_{n+1}$) به فرم زیر:

$$\tau_{n+1} = \sum_{k=n+1}^{N} k f_k(X_k) \prod_{j=n+1}^{k-1} (1 - f_j(X_j)) \tag{۴}$$

برای توابع اندازه پذیر $f_{n+1},\ldots,f_N:\mathbb{R}^d\to\{0,1\}$ با $f_N\equiv 1$ باشد. آن گاه یک تابع اندازه پذیر $f_n:\mathbb{R}^d\to\{0,1\}$ به صورتی که زمان توقف داده شده به فرم (۳) $\tau_n\in\mathcal{T}_n$ وجود دارد که شرط زیر را ارضا می کند:

$$\mathbb{E}[g(\tau_n,X_n)]\geq V_n-(V_{n+1}-\mathbb{E}[g(\tau_{n+1},X_{n+1})]), \tag{۵}$$

که در آن V_n و V_{n+1} مقادیر بهینه تعریف شده در (۲) می باشد.

اثبات. قرار دهید:

$$\epsilon=V_{n+1}-\mathbb{E}[g(\tau_{n+1},X_{n+1})]$$

و یک زمان توقف دلخواه مانند $\tau\in\mathcal{T}_n$ در نظر بگیرید. با توجه به لم (۱)، می دانیم که یک تابع اندازه پذیر مانند$h_n:\mathbb{R}^d\to\mathbb{R}$ وجود دارد به صورتی که $h_n(X_n)$ امید ریاضی شرطی $\mathbb{E}g(\tau_{n+1},X_{n+1}|X_n)$ می باشد. (توجه شود که تابع g، انتگرال پذیر است.)
علاوه بر این با توجه به فرم خاص τ_{n+1} در (۴)، داریم:

$$g(\tau_{n+1},X_{\tau_{n+1}})=\sum_{k=n+1}^{N}g(k,X_k)1_{\{\tau_{n+1=k}\}}$$
$$=\sum_{k=n+1}^{N}g(k,X_k)1_{\{f_k(X_k)\prod_{j=n+1}^{k-1}(1-f_j(X_j))=1\}}$$

که یک تابع اندازه پذیر از $X_{n+1},\ldots,X_N$ می باشد. (چرا که آن را می توان به صورت جمع توابع اندازه پذیر از $X_{n+1},\ldots,X_N$ نوشت.)
پس دیدیم که تابع $g(\tau_{n+1},X_{\tau_{n+1}})$ یک تابع اندازه پذیر از $X_{n+1},\ldots,X_N$ می باشد، پس با توجه به خاصیت مارکفی فرآیند $(X_n)_{n=0}^N$، داریم:

$$h_n(X_n)=\mathbb{E}[g(\tau_{n+1},X_{n+1}|X_n)]=\mathbb{E}[g(\tau_{n+1},X_{n+1}|\mathcal{F}_n)]$$

از آن جایی که پیشامد های

$$D=\{g(n,X_n)\geq h_n(X_n)\}\quad and\quad E=\{\tau=n\}$$

در $\mathcal{F}_n$ قرار دارند، $\tau_n=n1_D+\tau_{n+1}1_{D^c}$ متعلق به $\mathcal{T}_n$ و $\tilde{\tau}=\tau_{n+1}I_E+\tau I_{E^c}$ متعلق به $\mathcal{T}_{n+1}$ می باشند. بنابراین داریم:

$$\begin{aligned}\mathbb{E}[g\left(\tau_{n+1},X_{\tau_{n+1}}\right)] &=V_{n+1}-\epsilon && \text{با توجه به تعریف } \epsilon\\ &=\sup_{\tau\in\mathcal{T}_n}\mathbb{E}[g\left(\tau,X_\tau\right)]-\epsilon && \text{با توجه به تعریف } V_n\\ &\geq\mathbb{E}[g\left(\tilde{\tau},X_{\tilde{\tau}}\right)]-\epsilon\end{aligned}$$

و پس

$$\mathbb{E}[g(\tau_{n+1}, X_{\tau_{n+1}})] \geq \mathbb{E}[g(\tilde{\tau}, X_{\tilde{\tau}})] - \epsilon.$$

و لذا

$$\mathbb{E}[g(\tau_{n+1}, X_{\tau_{n+1}})1_{E^c}] \geq \mathbb{E}[g(\tilde{\tau}, X_{\tilde{\tau}})1_{E^c}] - \epsilon = \mathbb{E}[g(\tau, X_\tau)1_{E^c}] - \epsilon \qquad (۶)$$

توجه شود که $1_{E^c} = 1$ به این معنا می باشد که $\tilde{\tau} = \tau$.
در نهایت، به این می رسیم که:

$$\begin{aligned}
\mathbb{E}[g(\tau_n, X_{\tau_n})] &= \mathbb{E}[g(n, X_n)1_D + g(\tau_{n+1}, X_{\tau_{n+1}})1_{D^c}] && \text{با توجه به تعریف } \tau_{n+1} \\
&= \mathbb{E}[g(n, X_n)1_D + h_n(X_n)1_{D^c}] && \text{با توجه به تعریف } h_n(X_n) \\
&\geq \mathbb{E}[g(n, X_n)1_E + h_n(X_n)1_{E^c}] && \text{با توجه به } (\circledast) \\
&= \mathbb{E}[g(n, X_n)1_E + g(\tau_{n+1}, X_{\tau_{n+1}})1_{E^c}] && \text{با توجه به تعریف } h_n(X_n) \\
&\geq \mathbb{E}[g(n, X_n)1_E + g(\tau, X_\tau)1_{E^c}] - \epsilon && \text{با توجه به (۵)} \\
&= \mathbb{E}[g(\tau, X_\tau)] - \epsilon
\end{aligned}$$

$(\circledast)$: چرا که می دانیم اگر $g(n, X_n) \geq h_n(X_n)$، آن گاه $1_D = 1$ می باشد. از آن جایی که $\tau \in \mathcal{T}_n$ به دلخواه انتخاب شده بود، پس این نشان می دهد که $\mathbb{E}[g(\tau_n, X_{\tau_n})] \geq V_n - \epsilon$ می باشد و به این صورت، معادله ی (۵) برآورده شده است.
حال در نهایت، باید ثابت کنیم که τ_nای که یافته ایم، همانند (۳) است.
حال تابع $f_n : \mathbb{R}^d \to \{0, 1\}$ را به این صورت تعریف می کنیم:

$$f_n(x) \begin{cases} 1 & \text{اگر } g(n, x) \geq h_n(x) \\ 0 & \text{اگر } g(n, x) < h_n(x) \end{cases}$$

با توجه به تعریف پیشامد D داریم: $I_D = f_n(X_n)$. بنابراین داریم

$$\begin{aligned}
\tau_n &= n f_n(X_n) + \tau_{n+1}(1 - f_n(X_n)) && \text{با توجه به تعریف } \tau_n \\
&= \sum_{k=n}^{N} k f_k(X_k) \prod_{j=n}^{k-1} (1 - f_j(X_j)) && \text{با توجه به تعریف } \tau_{n+1} \text{ در صورت قضیه}
\end{aligned}$$

به همان چیزی که در نظر داشتیم، رسیدیم و پس اثبات به پایان می رسد.

□

توجه: قضیه ی (۱) نشان می دهد که زمان توقف داده شده در معادله(۳) را می توان در جهت محاسبه کردن V_n استفاده کرد. یعنی به وضوح، نابرابری(۵) معادل است با:

$$\sup_{\tau \in \mathcal{T}_{n+1}} \mathbb{E}g(\tau, X_\tau) - \mathbb{E}g\left(\tau_{n+1}, X_{\tau_{n+1}}\right) \geq \sup_{\tau \in \mathcal{T}_n} \mathbb{E}g(\tau, X_\tau) - g\left(\tau_n, X_{\tau_N}\right)$$

بنابراین از آن جایی که می دانیم $\tau_N = Nf\left(X_N\right)$، با یک استقراء بازگشتی می توان دید که τ_n، یک زمان توقف بهینه مناسب را به ما می دهد.

توجه: از آن جایی که $f_N \equiv 1$ می باشد، زمان توقف $\tau_N = f_N\left(X_N\right)$ یک زمان توقف بهینه در $\mathcal{T}_N$ می باشد. قضیه ی (۱) به صورت استقرایی، تواع اندازه پذیر $f_n : \mathbb{R}^d \to \{0,1\}$ را به ما می دهد به صورتی که برای همه $n \in \{0,1,\ldots,N-1\}$، زمان توقف داده شده ی τ_n توسط (۳)، در $\mathcal{T}_n$، نقطه ی توقف بهینه می باشد. به خصوص،

$$\tau = \sum_{n=1}^{N} nf_n\left(X_n\right)\prod_{j=0}^{n-1}\left(1-f_j\left(X_j\right)\right) \tag{۷}$$

یک زمان توقف بهینه برای مسئله توقف بهینه می باشد.

توجه: از قضیه ی (۱) این نیز به دست می اید که همه ی زمان های توقف متناظر با $V = \sup_{\tau\in\mathcal{T}}\mathbb{E}g\left(\tau, X_\tau\right)$ توسط معادله ی زیر به دست می آید:

$$\tau = \sum_{n=1}^{N} nf_n\left(X_n\right)\prod_{k=0}^{n-1}\left(1-f_k\left(X_k\right)\right)$$

توجه: در بسیاری از کاربرد ها، فرآیند مارکف X، از یک مقدار اولیه ی قطعی مانند $x_0 \in \mathbb{R}^d$ شروع می شود. پس تابع f_0 فقط با مقدار $f_0\left(x_0\right) \in \{0,1\}$ وارد ساختار (نمایش) (۷) می شود. یعنی در زمان ۰، تنها نیاز به یادگیری یک ثابت می باشد و نه کل تابع.

۳ معرفی یک شبکه عصبی

حال نشان دادیم که زمان توقف بهینه را می توان به صورت تابعی از دنباله ی تصمیمات توقف صفر و یکی، $\{f_n\}_{n=0}^{N}$، محاسبه کرد، نیاز دارم تا راهی برای تقریب زدن تابع نام برده بیابیم. روشی که می خواهیم از آن استفاده کنیم، توسط فلانی پیشنهاد شده است و با به کار گیری از یک شبکه ی عصبی می باشد. به طور خاص، دنباله ای از شبکه های عصبی به فرم : f^{θ_n} $\mathbb{R}^d \to \{0,1\}$ با پارامتر $\theta_n \in \mathbb{R}^q$ برای تقریب زدن f_n می سازیم. سپس، به سادگی می توانیم τ_{n+1} را با

$$\sum_{k=n}^{N} kf^{\theta_k}\left(X_k\right)\prod_{j=n}^{k-1}\left(1-f^{\theta_j}\left(X_j\right)\right) \tag{۸}$$

تقریب بزنیم. توجه شود که ما در گام n، به τ_{n+1} اشاره می کنیم نه τ_n. از آن جایی که به τ_{n+1} برای محاسبه ی ارزش ادامه دادن در زمان n نیاز داریم. ($V_{n+1} = \mathbb{E}\left[g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\right]$).

همان طور که در ادامه خواهیم دید، V_{n+1} نقشی کلیدی در تابع پادای که برای آموزش شبکه عصبی ایفا می کند.
پس اگر بخواهیم به صورت دقیق تر بیان کنیم:
روش عددی ما برای حل مسئله ی توقف بهینه بر پایه به صورت تکرار شونده (itereatively) تقریب زدن تصمیمات توقف بهینه $f_n : \mathbb{R}^d \to \{0,1\}, n = 0,1,\ldots,N-1$ به وسیله ی یک شبکه عصبی $f^\theta : \mathbb{R}^d \to \{0,1\}$ با پارامتر $\theta \in \mathbb{R}^q$ می باشد. این کار را به این صورت انجام میدهیم که تصمیم توقف نهایی را به صورت $f_N \equiv 1$ در نظر می گیریم و با اعمال استقرای بازگشتی (backward induction) شروع می کنیم. به طور دقیق تر، قرار دهید $n \in \{0,1,\ldots,N-1\}$ و فرض کنید مقادیر پارامتر ها $\theta_{n+1},\theta_{n+2},\ldots,\theta_N \in \mathbb{R}^q$ به صورتی در نظر گرفته شده اند که $f^{\theta_N} \equiv 1$ و زمان توقف

$$\tau_{n+1} = \sum_{m=n+1}^{N} m f^{\theta_m}\left(X_m\right) \prod_{j=n+1}^{m-1}\left(1 - f^{\theta_j}\left(X_j\right)\right)$$

یک امید ریاضی $\mathbb{E}[g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)]$ نزدیک به مقدار بهینه V_{n+1}. تولید می کند. اولین مشکلی که با آن مواجه هستیم، این است که خروجی تابع f^{θ_n}، فقط دارای مقادیری در $\{0,1\}$ می باشد. و در نتیجه با توجه به θ_n، پیوسته نمی باشد. این یک مشکل جدی می باشد چرا که در آموزش پارامترها برای یک شبکه عصبی، ما اغلب روش های بهینه سازی مبتنی بر گرادیان (gradient-based) را برای ماکسیمم کردن تابع پاداش (یا معادلاً مینیمم کردن ضرر) را به کار می بریم. همان طور که گفتیم، از آن جایی که f^θ مقادیری از $\{0,1\}$ اخذ می کند، پس نمی توانیم به صورت مستقیم از روش های بهینه سازی میتنی بر گرادیان استفاده کنیم. برای حل این مشکل، پارامتر ها را با یک شبکه عصبی چند لایه پیشخور آموزش می دهیم که خروجی های آن به صورت احتمال هایی که در بازه ی $(0,1)$ می باشند. سپس بعد از آموزش، می توانیم آن ها را به تصمیمات توقف صفر و یکی تبدیل کنیم. یعنی، برای $\theta \in \{\theta_0,\ldots,\theta_N\}$، $F^\theta : \mathbb{R}^d \to (0,1)$ را معرفی می کنیم که یک شبکه عصبی به فرم زیر می باشد:

$$F^\theta = \psi \circ a_I^\theta \circ \varphi_{q_{I-1}} \circ a_{I-1}^\theta \circ \cdots \circ \varphi_{q_1} \circ a_1^\theta$$

که در آن

- $I, q_1, q_2, \ldots, q_{I-1}$ اعداد صحیح مثبت هستند و عمق شبکه ی عصبی و تعداد نود های موجود در هر لایه ی نهان شبکه عصبی را نشان می دهند. (در صورت وجود لایه نهان)
- $a_1^\theta : \mathbb{R}^d \to \mathbb{R}^{q_1}, \ldots, a_{I-1}^\theta : \mathbb{R}^{q_{I-2}} \to \mathbb{R}^{q_{I-1}}$ و $a_I^\theta : \mathbb{R}^{q_{I-1}} \to \mathbb{R}$ توبع آفین (affine funtions) یا همان توبع تبدیلات خطی هستند.
- برای $j \in \mathbb{N}$ ، $\varphi_j : \mathbb{R}^j \to \mathbb{R}^j$ یک تابع فعال ساز ReLU می باشد که به صورت جزء به جزء component-wise اعمال می شود و به صورت زیر است:

$$\varphi_j\left(x_1,\ldots,x_j\right) = \left(x_1^+,\ldots,x_j^+\right)$$

- $\psi : \mathbb{R} \to (0,1)$ تابع لوجستیک استاندارد می باشد و به صورت زیر است:

$$\psi(x) = e^x / \left(1 + e^x\right) = 1 / \left(1 + e^{-x}\right)$$

اجزاء پارامتر $\theta \in \mathbb{R}^q$ از F^θ شامل ورودی هایی از ماتریس های $A_1 \in \mathbb{R}^{q_1 \times d}, \ldots, A_{I-1} \in \mathbb{R}^{q_{I-1} \times q_{I-2}}, A_I \in \mathbb{R}^{1 \times q_{I-1}}$ و بردار های $b_1 \in \mathbb{R}^{q_1}, \ldots, b_{I-1} \in \mathbb{R}^{q_{I-1}}, b_I \in \mathbb{R}$ می باشد که در ساختار توابع آفین که به صورت زیر هستند، می باشد:

$$a_i^\theta(x) = A_i x + b_i, \quad i = 1, \ldots, I.$$

پس بعد فضای پارامتر ها به صورت زیر می باشد:

$$q = \begin{cases} d+1 & \text{اگر } I = 1 \\ 1 + q_1 + \cdots + q_{I-1} + dq_1 + \cdots + q_{I-2}q_{I-1} + q_{I-1} & \text{اگر } I \geq 2 \end{cases}$$

و همچنین برای $x \in \mathbb{R}^d$ داده شده، $F^\theta(x)$، تابعی پیوسته و تقریباً همه جا در θ هموار می باشد. هدف ما تعیین کردن $\theta_n \in \mathbb{R}^q$ می باشد به صورتی که

$$\mathbb{E}\left[g\left(n, X_n\right) F^{\theta_n}\left(X_n\right) + g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1 - F^{\theta_n}\left(X_n\right)\right)\right]$$

مقداری نزدیک به $\sup_{\theta \in \mathbb{R}^q} \mathbb{E}\left[g\left(n, X_n\right) F^\theta\left(X_n\right) + g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1 - F^\theta\left(X_n\right)\right)\right]$ باشد. وقتی که به این هدف رسیدیم، تابع $f^{\theta_n} : \mathbb{R}^d \to \{0,1\}$ را به صورت زیر تعریف می کنیم:

$$f^{\theta_n} = 1_{[0,\infty)} \circ a_I^{\theta_n} \circ \varphi_{q_{I-1}} \circ a_{I-1}^{\theta_n} \circ \cdots \circ \varphi_{q_1} \circ a_1^{\theta_n}, \quad (۹)$$

که در آن، $1_{[0,\infty)} : \mathbb{R} \to \{0,1\}$ تابع مشخصه (indicator function) از $[0,\infty)$ می باشد. توجه شود که تنها تفاوتی که F^{θ_n} با f^{θ_n} دارد، غیرخطی بودن نهایی است.
توجه: توابع F^{θ_n} و f^{θ_n}:

$$F^\theta = \psi \circ a_I^\theta \circ \varphi_{q_{I-1}} \circ a_{I-1}^\theta \circ \cdots \circ \varphi_{q_1} \circ a_1^\theta$$
$$f^{\theta_n} = 1_{[0,\infty)} \circ a_I^{\theta_n} \circ \varphi_{q_{I-1}} \circ a_{I-1}^{\theta_n} \circ \cdots \circ \varphi_{q_1} \circ a_1^{\theta_n}$$

F^{θ_n} یک احتمال توقف در بازه ی $(0,1)$ تولید می کند در صورتی که f^{θ_n}، یک تصمیم توقف مشخص دقیق صفر یا یکی با توجه به این که f^{θ_n} دارای مقدار زیر $1/2$ باشد یا بالای آن، به ما می دهد.
حال که یک نسخه ی هموار از تصمیمات توقف تعریف کرده ایم و یک روش نیز، برای تبدیل خروجی های آن به تصمیمات توقف صفر و یکی معرفی کرده ایم، باید تابع پاداش (reward function) یا ضرر (loss function) برای میزان (درست) کردن (tune) پارامتر های θ_n تعریف کنیم. در هر گام، هدف ما ماکسیمم کردن امید ریاضی پاداش آینده می باشد. ما در زمان n

می دانیم که اگر توقف کنیم $(f_n(X_n) = 1)$و پاداشی به مقدار $g(n, X_n)$ دریافت می کنیم و اگر ادامه دهیم $(f_n(X_n) = 0)$ و پس از آن نیز به رفتار بهینه ی خود ادامه دهیم، در نهایت پاداشی به مقدار $g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)$ دریافت خواهیم کرد. (با توجه به قضیه ی (۱)) لذا در هر گام زمانی، به دنبال پیدا کردن تابع $f : \mathbb{R}^d \to \{0,1\}$ می باشیم که عبارت زیر را ماکسیمم کند:

$$\mathbb{E}\left[g(n, X_n) f(X_n) + g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)(1 - f(X_n))\right] \tag{۱۰}$$

۱.۳ قضیه (۲)

قضیه ای که در ادامه خواهیم دید، اثبات می کند که ما می توانیم f در معادله ی (۱۰)) را با f^θ جایگزین کنیم. این کار به ما این اجازه را می دهد که به جای پیدا کردن تابع بهینه $f : \mathbb{R}^d \to \{0,1\}$، معادله ی (۱۰)) را با توجه به $\theta \in \mathbb{R}^q$ ماکسیمم کنیم.

قضیه ۲. قرار دهید $n \in \{0, 1, \ldots, N-1\}$ و یک زمان توقف مانند $\tau_{n+1} \in \mathcal{T}_{n+1}$ را فیکس در نظر بگیرید. آن گاه، برای هر عمق $I \geq 2$ و ثابت $\varepsilon > 0$، اعداد صحیح مثبت $q_1, \ldots, q_{I-1}$ وجود دارند به صورتی که

$$\begin{aligned}&\sup_{\theta \in \mathbb{R}^q} \mathbb{E}\left[g(n, X_n) f^\theta(X_n) + g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1 - f^\theta(X_n)\right)\right] \\ &\geq \sup_{f \in \mathcal{D}} \mathbb{E}\left[g(n, X_n) f(X_n) + g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)(1 - f(X_n))\right] - \varepsilon.\end{aligned} \tag{۱۱}$$

که در آن، $\mathcal{D}$ مجموعه ی تمام توابع اندازه پذیر $f : \mathbb{R}^d \to \{0,1\}$ می باشد.

اثبات. ابتدا $\varepsilon > 0$ را فیکس می کنیم. در ابتدای مقاله اشاره شد که تابع g، تابعی انتگرال پذیر می باشد و پس در شرط انتگرال پذیری صدق می کند. همچنین، شرطی کمی قوی تر از انتگرال پذیری را نیز برای تابع g فرض کردیم و آن شرط به صورت زیر بود:

$$\mathbb{E}[g(n, X_n)^2] < \infty \quad \text{برای همه ی} \quad n \in \{0, 1, \ldots, N-1\} \tag{۱۲}$$

از شرط بالا نتیجه می شود که یک تابع اندازه پذیر مانند $\tilde{f} : \mathbb{R}^d \to \{0,1\}$ وجود دارد به صورتی که:

$$\begin{aligned}&\mathbb{E}\left[g(n, X_n) \tilde{f}(X_n) + g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1 - \tilde{f}(X_n)\right)\right] \\ &\geq \sup_{f \in \mathcal{D}} \mathbb{E}\left[g(n, X_n) f(X_n) + g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)(1 - f(X_n))\right] - \varepsilon/4.\end{aligned} \tag{۱۳}$$

$\tilde{f}$ را می توان به صورت $\tilde{f} = 1_A$ برای مجموعه برل (Borel set) A که به صورت $A = \left\{x \in \mathbb{R}^d : \tilde{f}(x) = 1\right\}$ می باشد، نوشت. علاوه بر این، با توجه به (۱۱))،

$$B \mapsto \mathbb{E}\left[|g(n, X_n)| 1_B(X_n)\right] \quad \text{و} \quad B \mapsto \mathbb{E}\left[\left|g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\right| 1_B(X_n)\right]$$

اندازه برل متناهی روی $\mathbb{R}^d$ تعریف می کند. از آن جا که هر اندازه برل متناهی روی $\mathbb{R}^d$ $tight$ می باشد، یک زیرمجموعه فشرده که می تواند تهی نیز باشد، مانند $K \subseteq A$ وجود دارد به صورتی که:

$$\begin{aligned}
&\mathbb{E}\left[g\left(n, X_n\right) 1_K\left(X_n\right)+g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1-1_K\left(X_n\right)\right)\right] \\
&\geq \mathbb{E}\left[g\left(n, X_n\right) \tilde{f}\left(X_n\right)+g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1-\tilde{f}\left(X_n\right)\right)\right]-\varepsilon / 4.
\end{aligned} \tag{۱۴}$$

$\rho_K : \mathbb{R}^d \to [0, \infty]$ را یک تابع فاصله که به صورت $\rho_K(x) = \inf_{y \in K} \|x-y\|_2$ می باشد، در نظر بگیرید. آن گاه

$$k_j(x)=\max \left\{1-j \rho_K(x),-1\right\}, \quad j \in \mathbb{N},$$

دنباله ای از توابع پیوسته $k_j : \mathbb{R}^d \to [-1, 1]$ را تعیین می کند که به صورت نقطه ای به $1_K - 1_{K^c}$ همگرا می باشند. حال با توجه به قضیه ی همگرایی تسلطی لبگ (Lebesgue's dominated convergence) می دانیم که یک $j \in \mathbb{N}$ وجود دارد به صورتی که:

$$\begin{aligned}
&\mathbb{E}\left[g\left(n, X_n\right) 1_{\left\{k_j\left(X_n\right) \geq 0\right\}}+g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1-1_{\left\{k_j\left(X_n\right) \geq 0\right\}}\right)\right] \\
&\geq \mathbb{E}\left[g\left(n, X_n\right) 1_K\left(X_n\right)+g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1-1_K\left(X_n\right)\right)\right]-\varepsilon / 4.
\end{aligned} \tag{۱۵}$$

با استفاده از قضیه ی (۱) (Leshno) می بینیم که k_j را می توان به صورت یکنواخت (uniformly) روی مجموعه های فشرده با توابعی به فرم پایین تقریب زد:

$$\sum_{i=1}^r\left(v_i^T x+c_i\right)^{+}-\sum_{i=1}^s\left(w_i^T x+d_i\right)^{+} \tag{۱۶}$$

برای $r, s \in \mathbb{N}, v_1, \ldots, v_r, w_1, \ldots, w_s \in \mathbb{R}^d$ و $c_1, \ldots, c_r, d_1, \ldots, d_s \in \mathbb{R}$. پس تابعی مانند $h : \mathbb{R}^d \to \mathbb{R}$ وجود دارد که قابل بیان به صورت (۱۶) می باشد به صورتی که:

$$\begin{aligned}
&\mathbb{E}\left[g\left(n, X_n\right) 1_{\left\{h\left(X_n\right) \geq 0\right\}}+g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1-1_{\left\{h\left(X_n\right) \geq 0\right\}}\right)\right] \\
&\geq \mathbb{E}\left[g\left(n, X_n\right) 1_{\left\{k_j\left(X_n\right) \geq 0\right\}}+g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1-1_{\left\{k_j\left(X_n\right) \geq 0\right\}}\right)\right]-\varepsilon / 4.
\end{aligned} \tag{۱۷}$$

حال توجه شود که برای هر عدد صحیح $I \geq 2$، ترکیب نگاشت $1_{[0,\infty)} \circ h$ را می توان به عنوان یک شبکه عصبی f^θ به فرم (۹) با عمق I برای اعداد صحیح مناسب $q_1, \ldots, q_{I-1}$ و مقدار پارامتر $\theta \in \mathbb{R}^q$ نوشت. بنابراین، با توجه به (۱۳)، (۱۴)، (۱۵) و (۱۷) به دست می آید که:

$$\begin{aligned}
&\mathbb{E}\left[g\left(n, X_n\right) f^\theta\left(X_n\right)+g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1-f^\theta\left(X_n\right)\right)\right] \\
&\geq \sup_{f \in \mathcal{D}} \mathbb{E}\left[g\left(n, X_n\right) f\left(X_n\right)+g\left(\tau_{n+1}, X_{\tau_{n+1}}\right)\left(1-f\left(X_n\right)\right)\right]-\varepsilon,
\end{aligned}$$

و پس اثبات به پایان رسید. $\square$

ما همواره، $\theta_N \in \mathbb{R}^q$ به گونه ای انتخاب می کنیم که $f^{\theta_N} \equiv 1$ باشد. (به راحتی می توان دید که این امر امکان پذیر است.) سپس زمات توقف بهینه کاندید مان

$$\tau^{\Theta} = \sum_{n=1}^{N} n f^{\theta_n}\left(X_n\right) \prod_{j=0}^{n-1} \left(1 - f^{\theta_j}\left(X_j\right)\right) \tag{۱۸}$$

که با بردار $\Theta = \left(\theta_0, \theta_1, \ldots, \theta_{N-1}\right) \in \mathbb{R}^{Nq}$ مشخص شده است. نتیجه ی زیر، نتیجه ی مستقیم از قضیه ی (۱) و (۲) می باشد:

نتیجه ۱. برای یک مسئله ی توقف بهینه به فرم مسئله اصلی، یک عمق $I \geq 2$ و ثابت $\varepsilon > 0$، اعداد صحیح مثبت $q_1, \ldots, q_{I-1}$ و بردار $\Theta \in \mathbb{R}^{Nq}$ به گونه ای وجود دارند که زمان توقف متناظر با (۱۸) که در شرط $\mathbb{E} g\left(\tau^{\Theta}, X_{\tau^{\Theta}}\right) \geq \sup_{\tau \in \mathcal{T}} \mathbb{E} g\left(\tau, X_\tau\right) - \varepsilon$ صدق می کند.

۲.۳ کران ها

در این بخش به پیدا کردن کران بالا و کران پایین برای مقدار بهینه $V_0 = \sup_{\tau \in \mathcal{T}} \mathbb{E} g\left(\tau, X_\tau\right)$ می پردازیم.

۱.۲.۳ کران پایین

وقتی که تصمیمات توقف f^{θ_n} آموزش داده شده اند، زمان توقف τ^{Θ} که توسط معادله ی (۱۸) داده شده است، یک کران پایین $L = \mathbb{E} g\left(\tau^{\Theta}, X_{\tau^{\Theta}}\right)$ برای مقدار بهینه $V_0 = \sup_{\tau \in \mathcal{T}} \mathbb{E} g\left(\tau, X_\tau\right)$ به ما می دهند. برای تخمین زدن آن، ما یک مجموعه ی جدید از وقایع مستقل $\left(y_n^k\right)_{n=0}^N, k = 1, 2, \ldots, K_L$، از $\left(X_n\right)_{n=0}^N$ را شبیه سازی می کنیم. برای یک تابع اندازه پذیر $l : \mathbb{R}^{dN} \to \{0, 1, \ldots, N\}$، τ^{Θ} به فرم $\tau^{\Theta} = l\left(X_0, \ldots, X_{N-1}\right)$ می باشد. قرار دهید $l^k = l\left(y_0^k, \ldots, y_{N-1}^k\right)$. تقریب مونته-کارلو (Monte Carlo)

$$\hat{L} = \frac{1}{K_L} \sum_{k=1}^{K_L} g\left(l^k, y_{l^k}^k\right)$$

یک تخمین بی طرفانه و منصفانه از کران پایین L می دهد، و با توجه به قانون اعداد بزرگ، هنگامی که $K_L \to \infty$، $\hat{L}$ به L همگرا می شود. فرض می کنیم که نمونه های $\left(y_n^k\right)_{n=0}^N, k = 1, \ldots, K_L$، به صورت مستقل از وقایع $\left(x_n^k\right)_{n=0}^N, k = 1, \ldots, K$ اخذ شده اند که در آموزش دادن تصمیمات توقف از آن ها استفاده کرده ایم.)

۲.۲.۳ کران بالا

(Snell envelope) یا پوشش (Snell) از فرآیند پاداش $\left(g\left(n, X_n\right)\right)_{n=0}^N$، کوچکترین سوپرمارتینگل (supermartingale) با توجه به $\left(\mathcal{F}_n\right)_{n=0}^N$ می باشد که بر $\left(g\left(n, X_n\right)\right)_{n=0}^N$ تسلط دارد و به

صورت زیر می باشد:

$$H_n = \mathrm{ess}_{\sup_{\tau\in\mathcal{T}_n}} \mathbb{E}\left[g(\tau) \mid \mathcal{F}_n\right], \quad n = 0, 1, \ldots, N$$

تجزیه ی (Doob-Mayer) آن به صورت زیر است:

$$H_n = H_0 + M_n^H - A_n^H$$

که در آن، M^H، $(\mathcal{F}_n) - martingale$ است و به صورت زیر می باشد:

$$M_0^H = 0 \quad \text{و} \quad M_n^H - M_{n-1}^H = H_n - \mathbb{E}\left[H_n \mid \mathcal{F}_{n-1}\right], \quad n = 1, \ldots, N,$$

و A^H، یک فرآیند $(\mathcal{F}_n) - predictable$ غیر نزولی می باشد که به صورت زیر است:

$$A_0^H = 0 \quad \text{و} \quad A_n^H - A_{n-1}^H = H_{n-1} - \mathbb{E}\left[H_n \mid \mathcal{F}_{n-1}\right], \quad n = 1, \ldots, N.$$

تخمین ما برای کران بالا برای مقدار بهینه V_0، بر اساس ورژنی از فرمول بندی دوگان برای مسائل توقف بهینه می باشد که توسط Rogers (2002) و Haugh و Kogan (2004) معرفی شده است.

قضیه ۳. فرض کنید $(\varepsilon_n)_{n=0}^N$ یک دنباله ای متغیر های تصادفی اتگرال پذیر روی فضای احتمالاتی $(\Omega, \mathcal{F}, \mathbb{P})$ باشد. آن گاه داریم:

$$V_0 \geq \mathbb{E}\left[\max_{0\leq n\leq N}\left(g\left(n, X_n\right) - M_n^H - \varepsilon_n\right)\right] + \mathbb{E}\left[\min_{0\leq n\leq N}\left(A_n^H + \varepsilon_n\right)\right]. \qquad (۱۹)$$

و علاوه بر این، اگر برای همه ی $n \in \{0, 1, \ldots, N\}$ داشته باشیم $\mathbb{E}\left[\varepsilon_n \mid \mathcal{F}_n\right] = 0$، یکی دارای

$$V_0 \leq \mathbb{E}\left[\max_{0\leq n\leq N}\left(g\left(n, X_n\right) - M_n - \varepsilon_n\right)\right] \qquad (۲۰)$$

می باشد برای هر $(\mathcal{F}_n) - martingale$ $(M_n)_{n=0}^N$ که از 0 شروع می شود.

اثبات. ابتدا توجه کنید که:

$$\begin{aligned}\mathbb{E}\left[\max_{0\leq n\leq N}\left(g\left(n, X_n\right) - M_n^H - \varepsilon_n\right)\right] &\leq \mathbb{E}\left[\max_{0\leq n\leq N}\left(H_n - M_n^H - \varepsilon_n\right)\right] \\ &= \mathbb{E}\left[\max_{0\leq n\leq N}\left(H_0 - A_n^H - \varepsilon_n\right)\right] = V_0 - \mathbb{E}\left[\min_{0\leq n\leq N}\left(A_n^H + \varepsilon_n\right)\right],\end{aligned}$$

که (۱۹) نشان می دهد.
حال فرض می کنیم که برای تمام $n \in \{0, 1, \ldots, N\}$، $\mathbb{E}\left[\varepsilon_n \mid \mathcal{F}_n\right] = 0$ برقرار است. τ را نیز یک $X - stoppingtime$ در نظر بگیرید. آن گاه

$$\mathbb{E}\varepsilon_\tau = \mathbb{E}\left[\sum_{n=0}^N 1_{\{\tau=n\}}\varepsilon_n\right] = \mathbb{E}\left[\sum_{n=0}^N 1_{\{\tau=n\}}\mathbb{E}\left[\varepsilon_n \mid \mathcal{F}_n\right]\right] = 0.$$

پس با توجه به قضیه ی توقف بهینه داریم:

$$\mathbb{E}g(\tau, X_\tau) = \mathbb{E}\left[g(\tau, X_\tau) - M_\tau - \varepsilon_\tau\right] \leq \mathbb{E}\left[\max_{0\leq n\leq N}\left(g(n, X_n) - M_n - \varepsilon_n\right)\right]$$

برای هر $(\mathcal{F}_n) - martingale$ $(M_n)_{n=0}^N$ که از 0 شروع می شود. از آن جایی که $V_0 = \sup_{\tau\in\mathcal{T}}\mathbb{E}g(\tau, X_\tau)$، پس (۲۰) را می رساند. $\square$

پس برای هر $(\mathcal{F}_n) - martingale$ با شروع از 0 و هر دنباله ای از ترم های انتگرال پذیر خطا $(\varepsilon_n)_{n=0}^N$، که برای تمام nها در شرط $\mathbb{E}\left[\varepsilon_n \mid \mathcal{F}_n\right] = 0$ صدق می کنند، سمت راست (۲۰)، ی کران بالا برای V_0 به ما می دهد. (توجه شود که برای این که سمت راست (۲۰) یک کران بالا قابل قبول به ما بدهد، کافی است که برای تمام nها، داشته باشیم $\mathbb{E}\left[\varepsilon_n \mid \mathcal{F}_n\right] = 0$. به ویژه، توجه شود که $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_N$، می توانند هر ساختار وابستگی دلخواهی داشته باشند.) و با توجه به (۱۹)، اگر $M = M^H$ و $\varepsilon \equiv 0$ باشد، این کران بالا، یک کران بالا سفت و محکم (tight) می باشد. پس ما سعی می کنیم که تا با استفاده از کاندید توقف بهینه مان τ^Θ، یک مارتینگل نزدیک به M^H بسازیم. هر چقدر که τ^Θ نزدیکتر به زمان توقف باشد، فرآیند ارزش بهتر می باشد:

$$H_n^\Theta = \mathbb{E}\left[g\left(\tau_n^\Theta, X_{\tau_n^\Theta}\right) \mid \mathcal{F}_n\right], \quad n = 0, 1, \ldots, N$$

متناظر با:

$$\tau_n^\Theta = \sum_{m=n}^{N} m f^{\theta_m}(X_m) \prod_{j=n}^{m-1}\left(1 - f^{\theta_j}(X_j)\right), \quad n = 0, 1, \ldots, N,$$

پوشش (Snell) $(H_n)_{n=0}^N$ را تقریب می زند. قسمت مارتینگل $\left(H_n^\Theta\right)_{n=0}^N$، توسط $M_0^\Theta = 0$ داده شده است و

$$\begin{aligned} M_n^\Theta - M_{n-1}^\Theta &= H_n^\Theta - \mathbb{E}\left[H_n^\Theta \mid \mathcal{F}_{n-1}\right] \\ &= f^{\theta_n}(X_n)\, g(n, X_n) + \left(1 - f^{\theta_n}(X_n)\right) C_n^\Theta - C_{n-1}^\Theta, n \geq 1, \end{aligned} \quad (۲۱)$$

و برای مقادیر ادامه دادن

$$C_n^\Theta = \mathbb{E}\left[g\left(\tau_{n+1}^\Theta, X_{\tau_{n+1}^\Theta}\right) \mid \mathcal{F}_n\right] = \mathbb{E}\left[g\left(\tau_{n+1}^\Theta, X_{\tau_{n+1}^\Theta}\right) \mid X_n\right], n = 0, 1, \ldots, N-1.$$

توجه شود که C_N^Θ نباید از قبل تعیین شده باشد. این در (۲۱) برای $n = N$ رسماً ظاهر می شود. ولی $\left(1 - f^{\theta_N}(X_N)\right)$ همواره 0 نمی باشد. برای تخمین M^Θ، مجموعه ی سومی از وقایع مستقل $\left(z_n^k\right)_{n=0}^N, k = 1, 2, \ldots, K_U$ از $(X_n)_{n=0}^N$ تولید می کنیم. به علاوه، برای هر z_n^k، J تا ادامه ی مسیر های $\tilde{z}_{n+1}^{k,j}, \ldots, \tilde{z}_N^{k,j}, j = 1, \ldots, J$ را شبیه سازی می کنیم که به صورت

مشروط از یک دیگر و از $z_{n+1}^k, \ldots, z_N^k$ مستقل هستند. فرض کنید که ارزش τ_{n+1}^{Θ} را در طی $\tilde{z}_{n+1}^{k,j}, \ldots, \tilde{z}_N^{k,j}$ با $\tau_{n+1}^{k,j}$ نمایش دهیم. تخمین زدن ارزش ادامه به صورت زیر

$$C_n^k = \frac{1}{J} \sum_{j=1}^{J} g\left(\tau_{n+1}^{k,j}, \tilde{z}_{\tau_{n+1}^{k,j}}^{k,j}\right), \quad n = 0, 1, \ldots, N-1,$$

به ما یک تخمین نویزی (noisy) به صورت زیر:

$$\Delta M_n^k = f^{\theta_n}\left(z_n^k\right) g\left(n, z_n^k\right) + \left(1 - f^{\theta_n}\left(z_n^k\right)\right) C_n^k - C_{n-1}^k$$

از افزاش های $M_n^{\Theta} - M_{n-1}^{\Theta}$ در طی kامین مسیر شبیه سازی شده $z_0^k, \ldots, z_N^k$ به ما می دهد. پس:

$$M_n^k = \begin{cases} 0 & \text{اگر } n = 0 \\ \sum_{m=1}^{n} \Delta M_m^k & \text{اگر } n \geq 1 \end{cases}$$

را می توان به صورت وقایعی از $M_n^{\Theta} + \varepsilon_n$ برای تخمین خطاهای ε_n با انحراف های استاندارد متناسب با $1/\sqrt{J}$ به صورتی که برای تمام nها، $\mathbb{E}\left[\varepsilon_n \mid \mathcal{F}_n\right] = 0$ برقرار باشد، دید. بنابراین،

$$\hat{U} = \frac{1}{K_U} \sum_{k=1}^{K_U} \max_{0 \leq n \leq N} \left(g\left(n, z_n^k\right) - M_n^k\right),$$

یک تخمین بی طرف و منصفانه(unbiased) از کران بالا

$$U = \mathbb{E}\left[\max_{0 \leq n \leq N} \left(g\left(n, X_n\right) - M_n^{\Theta} - \varepsilon_n\right)\right],$$

می باشد که، با توجه به قانون اعداد بزرگ، برای $K_U \to \infty$ به U همگرا می باشد.

۳.۲.۳ برآورد نقطه ای و بازه ی اطمینان

برآورد نقطه ای ما از V_0، میانگین به صورت زیر می باشد:

$$\frac{\hat{L} + \hat{U}}{2}$$

برای بدست آوردن بازه ی اطمینان، فرض می کنیم که $g\left(n, X_n\right)$ برای تمام nها، مربع-انتگرال پذیر (square-integrable) باشد. آن گاه:

$$g\left(\tau^{\theta}, X_{\tau^{\Theta}}\right) \quad \text{و} \quad \max_{0 \leq n \leq N} \left(g\left(n, X_n\right) - M_n^{\Theta} - \varepsilon_n\right)$$

این ها نیزمربع-انتگرال پذیر هستند. از قضیه ی حد مرکزی به دست می آید که برای K_L بزرگ، $\hat{L}$ به صورت تقریبی دارای توزیع نرمال با میانگین L و واریانس $\hat{\sigma}_L^2/K_L$ می باشد برای:

$$\hat{\sigma}_L^2 = \frac{1}{K_L - 1} \sum_{k=1}^{K_L} \left(g\left(l^k, y_{l^k}^k\right) - \hat{L} \right)^2$$

پس، برای هر $\alpha \in (0, 1]$:

$$\left[\hat{L} - z_{\alpha/2} \frac{\hat{\sigma}_L}{\sqrt{K_L}}, \infty\right)$$

یک بازه ی اطمینان مجانبی معتبر $1 - \alpha/2$ برای L می باشد، به صورتی که $z_{\alpha/2}$، $1 - \alpha/2$ چندک (quantile) از توزیع نرمال استاندارد می باشد. به طور مشابه:

$$\left(-\infty, \hat{U} + z_{\alpha/2} \frac{\hat{\sigma}_U}{\sqrt{K_U}}\right] \quad \text{با} \quad \hat{\sigma}_U^2 = \frac{1}{K_U - 1} \sum_{k=1}^{K_U} \left(\max_{0 \le n \le N} \left(g\left(n, z_n^k\right) - M_n^k \right) - \hat{U} \right)^2$$

یک بازه ی اطمینان مجانبی معتبر $1 - \alpha/2$ برای U می باشد. نتیجه می شود که برای هر ثابت $\varepsilon > 0$، تا زمانی که K_L و K_U به اندازه ی کافی بزرگ باشند، داریم:

$$\begin{aligned} &\mathbb{P}\left[V_0 < \hat{L} - z_{\alpha/2} \frac{\hat{\sigma}_L}{\sqrt{K_L}} \text{ یا } V_0 > \hat{U} + z_{\alpha/2} \frac{\hat{\sigma}_U}{\sqrt{K_U}}\right] \\ &\le \mathbb{P}\left[L < \hat{L} - z_{\alpha/2} \frac{\hat{\sigma}_L}{\sqrt{K_L}}\right] + \mathbb{P}\left[U > \hat{U} + z_{\alpha/2} \frac{\hat{\sigma}_U}{\sqrt{K_U}}\right] \le \alpha + \varepsilon \end{aligned}$$

به ویژه:

$$\left[\hat{L} - z_{\alpha/2} \frac{\hat{\sigma}_L}{\sqrt{K_L}}, \hat{U} + z_{\alpha/2} \frac{\hat{\sigma}_U}{\sqrt{K_U}}\right]$$

یک بازه ی اطمینان مجانبی معتبر $1 - \alpha$ برای V_0 می باشد. (بازه ی اطمینان، تنها در زمانی که اندازه ی نمونه به اندازه ی کافی بزرگ باشد، معتبر است. بازه ی اطمینان مجانبی نیز به همین صورت می باشد و پس بازه ی اطمینان مجانبی، تنها در نمونه های با اندازه های بسیار بزرگ کاربرد دارد. ایده ی اصلی بازه اطمینان مجانبی، بر اساس قضیه ی حد مرکزی می باشد، که در نمونه با اندازه های بزرگ، توزیع نمونه به طور تقریبی به توزیع نرمال همگرا می شود. بازه ی اطمینان مجانبی برای زمانی استفاده می شود که ما توزیع اولیه ی داده ها را در دست نداریم ولی با توجه به قضیه ی حد مرکزی، می دانیم که اگر اندازه نمونه به حد کافی بزرگ باشد، توزیع آن به توزیع نرمال میل می کند.)

۳.۳ شبیه سازی حرکت براونی هندسی

حرکت براونی هندسی (Geometric Brownian Motion) یا به اختصار(GBM)، در واقع یک فرآیند مارکف است. به این معنا که قیمت سهام به صورت یک قدم زدن تصادفی می باشد و

یکی از فرضیه های اساسی مدل بلک–شولز (Black-Sholes Model) می باشد. پس لازم است که مسیر های (GBM) که در آموزش و تست کردن مدل از آن ها استفاده می کنیم، شبیه سازی کنیم. فرمول (GBM) به صورت زیر می باشد:

$$\frac{\Delta S}{S} = \mu \Delta t + \sigma \epsilon \sqrt{\Delta t}$$

که در آن:

$S =$ قیمت سهام

$\Delta S =$ تغییر در قیمت سهام

$\mu =$ امید ریاضی بازده یا دیریفت (dfirf)

$\sigma =$ انحراف استاندارد بازده ها یا دیفیوژن (diffusion)

$\epsilon =$متغیر تصادفی

$\Delta t =$ بازه ی زمانی سپری شده

برای شبیه سازی یک مسیر از حرکت براونی در روش جلورونده، می توانیم از خاصیت استقلال و پایا افزایش های گاوسی استفاده کنیم و یا پل براونی Brownian bridge در یک شبیه سازی عقب رونده استفاده کنیم. یک پل براونی X، یک فرآیند پیوسته ی گاوسی با $X_0 = X_1 = 0$ می باشد و به معنی:

$$\mathbb{E}[(X_t)] = 0 \quad \text{برای} \; t \in [0,1]$$

و با کوواریانس:

$$\operatorname{cov}(X_s, X_t) = \min\{s,t\} - st \;\square\square\square\; s,t \in [0,1]$$

و سپس، از متود چولسکی (Cholesky's method)، برای تعمیم شبیه سازی از یک بعد به n بعد استفاده می کنیم. عکس قسمتی از کد شبیه سازی در زیر آورده شده است. (شکل (۴))

```
def stock_sim_path(S, alpha, delta, sigma, T, N, n):
    """Simulates geometric Brownian motion."""
    h = T/n
    # uncomment below for deterministic trend. or, can pass it in
                                        as alpha as an array
    alpha = alpha # + np.linspace(0, 0.1, 500).reshape((n,N))
    mean = (alpha - delta - .5*sigma**2)*h
    vol = sigma * h**.5
    return S*np.exp((mean + vol*np.random.randn(n,N)).cumsum(axis =
                                         0))
```

شکل ۴: قسمتی از کد شبیه سازی (GBM)

۴.۳ مثال ها

۱.۴.۳ Bermudan Max-Call Option

Bermudan Max-Call Optionای که در زمان T اجرا می شود، قرارداد اختیاری می باشد که روی d دارایی نوشته می شود، $\{X^i\}_{i=1}^d$، که به دارانده ی آن این اجازه را می دهد تا یکی از این دارایی ها را با قیمت توافقی K در هر مقطع زمانی در $0 = t_0 < t_1 < \ldots < t_N = T$ خریداری کند. ما فرض می کنیم که در یک بازار با مدل بلک–شولز هستیم و پس قیمت دارایی ها را با حرکت براونی زیر به دست می آیند:

$$X_t^i = (r - \gamma_i)\, dt + \sigma_i dW_i^{ti}$$

که در آن

- x_0^i، قسمت سهام در زمان 0 می باشد.
- $r \in [0, \infty)$ نرخ بهره بدون ریسک می باشد.
- $\gamma_i \in [0, \infty)$ عملکرد سود
- $\sigma_i \in [0, \infty)$ نوسلن پذیری دارایی
- W_i^t، که iامین درایه از حرکت براونی d بعدی می باشد.

تابع عایدی آن در زمان t به صورت زیر می باشد:

$$\sup_{\tau} \mathbb{E}\left[e^{-r\tau}\left(\max_{1\le i\le d} S_\tau^i - K\right)^+\right],$$

که در آن، سوپریمم روی تمام $S - stoppingtime$ها که مقادیر آن ها $t_0, t_1, \ldots, t_N$ می باشد. قسمتی از کد شبکه ی عصبی در ادامه آورده شده است.

```
class NeuralNet(torch.nn.Module):
    def __init__(self, d, q1, q2):
        super(NeuralNet, self).__init__()
        self.a1 = nn.Linear(d, q1)
        self.relu = nn.ReLU()
        self.a2 = nn.Linear(q1, q2)
        self.a3 = nn.Linear(q2, 1)
        self.sigmoid=nn.Sigmoid()

    def forward(self, x):
        out = self.a1(x)
        out = self.relu(out)
        out = self.a2(out)
        out = self.relu(out)
        out = self.a3(out)
        out = self.sigmoid(out)

        return out
```

شکل ۵: قسمتی از کد ساخت شبکه عصبی

```
def loss(y_pred,s, x, n, tau):
    r_n=torch.zeros((s.M))
    for m in range(0,s.M):

        r_n[m]=-s.g(n,m,x)*y_pred[m] - s.g(tau[m],m,x)*(1-y_pred[m]
                                        )

    return(r_n.mean())
```

شکل ۶: .تابع ضرر که می خواهیم آن را مینیمم کنیم

```
def NN(n,x,s, tau_n_plus_1):
    epochs=50
    model=NeuralNet(s.d,s.d+40,s.d+40)
    optimizer = torch.optim.Adam(model.parameters(), lr = 0.0001)

    for epoch in range(epochs):
        F = model.forward(X[n])
        optimizer.zero_grad()
        criterion = loss(F,S,X,n,tau_n_plus_1)
        criterion.backward()
        optimizer.step()

    return F,model
```

شکل ۷: کد آموزش داده ها

۲.۴.۳ Asian Max-call Option

همان طور که در اول گزارش اشاره شد، در قرارداد های اختیار آسیایی، بر خلاف قرار داد های اروپایی و آمریکایی، عایدی تنها وابسته قیمت دارایی پایه در یک مقطع زمانی خاص نمی باشد.
دو نوع قرارداد اختیار آسیایی داریم:
۱) (fixed strike): در این نوع آپشن، قیمت میانگین به جای قیمت دارایی پایه استفاده می شود.
۲) (fixed price): در این نوع آپشن، میانگین قیمت به جای قیمت توافقی استفاده می شود.
عایدی قرارداد اختیار آسیایی به صورت زبر محاسبه می شود:

$$C(T) = \max(A(0,T) - K, 0)$$

که در آن، $A(0,\ T)$، نمایانگر میانگین قیمت دارایی پایه در بازه ی زمانی 0 تا T می باشد.
قسمت هایی از کد در ادامه آورده شده است.

```
def g(self,n,m,X):
        max1=torch.max(X[int(n),m,:].float()-self.K)

        return np.exp(-self.r*(self.T/self.N)*n)*torch.max(max1,
                                                    torch.tensor([0.0]))
```

```python
class stock:
    def __init__(self, T, K, sigma, delta, So, r, N, M, d):
        self.T = T                      # ending time step
        self.K=K                        # price purchased
        self.sigma=sigma *np.ones(d)    # asset volatility
        self.delta=delta                # dividend yield
        self.So=So*np.ones(d)           # price at time 0
        self.r=r                        # marker risk free return
        self.N=N                        # number of time steps
        self.M=M                        # number of sample paths
        self.d=d                        # number of assets
        self.a=a                        # value of stock
        self.i=0                        # number of stopping

    def g(self,n,m,X):
        self.a = self.a + X
        self.i = self.i + 1
        Y = self.a/self.i

        max1=torch.max(Y[int(n),m,:].float()-self.K)

        return np.exp(-self.r*(self.T/self.N)*n)*torch.max(max1,
                                                    torch.tensor([0.0]))
```

شکل ۸: .تابع ضرر که می خواهیم آن را مینیمم کنیم

```python
def NN(n,x,s, tau_n_plus_1):
    epochs=50
    model=NeuralNet(s.d,s.d+40,s.d+40)
    optimizer = torch.optim.Adam(model.parameters(), lr = 0.0001)

    for epoch in range(epochs):
        F = model.forward(X[n])
        optimizer.zero_grad()
        criterion = loss(F,S,X,n,tau_n_plus_1)
        criterion.backward()
        optimizer.step()

    return F,model
```

۴ منابع

[1] Becker S, Cheridito P, Jentzen A. Deep optimal stopping[J]. Journal of Machine Learning Research, 2019, 20: 74.

[2] Becker S, Cheridito P, Jentzen A. Pricing and hedging American-style options with deep learning [J]. Journal of Risk and Financial Management, 2020, 13(7): 158 .

[3] Cheridito P, Jentzen A, Rossmannek F. Efficient approximation of high-dimensional functions with deep neural networks[J].

[4] Longstaff F A, Schwartz E S. Valuing American options by simulation: a simple least-squares approach [J]. The review of financial studies, 2001, 14(1): 113-147.

[5] Karl S, Simulating Brownian motion and geometric Brownian motion