שיטות נומריות

שיעור 1

מבנה הקורס 1. ייצור משתנים מקריים מהתפלגויות שונות (דגימה).

2. סימולציות מונטה קרלו.

3. תמחור אופציות אמריקאיות (נושא פחות מרכזי).

שיטות נומריות שיטה נומרית נותנת פתרונות לבעיות שאין להם פתרון מדויק או שהפתרון המדויק הוא יקר. במקרה שהפיתרון המדויק יקר, השיטה הנומרית היא פיתרון זול יותר המגיע על חשבון דיוק הפיתרון.

1. נתונה בעיה מציאותית־ למשל תמחור אופציה מורכבת.

2. בונים מודל לאותה בעיה (ניסוח מודל מתמטי). המודל נבנה תחת הנחות מסוימות. במקרה שלנו, לרוב המודל יהיה אוסף משוואות דיפרנציאליות סטוכססטיות $dx_t = ..dt + ...dW_t$ אחרי ניסוח המודל, נרצה להסיק ממנו מסקנות.

3. בשלב זה מחפשים פיתרון (נוסחה) למודל. הנוסחה יכולה להיות נוסחה מפורשת (כמו בלק שולס), אבל לרוב נקבל ביטוי שלא נוכל לחשב אלא רק לקרב.

אם אין פתרון מפורש, נפעיל שיטה נומרית ונמצא פתרון מקורב.

4. יישום השיטה נומרית (לרוב בעזרת מחשב).

דוגמא (שיטת נקודת השבת) נתונה המשוואה הבאה $x^2-3x+2=\frac{1}{12}sinx$. נרצה למצוא למשוואה פיתרון. נשים לב שהתחלנו במקרה זה בשלב השלישי, כאשר המודל המתמטי הוא משוואה. נשים לב שלא ניתן לחלץ את x ולכן נפעיל שיטה נומרית.

ננחש x מתאים, נציב באגף ימין ונחשב את $sin(x)$.

ננחש $x_1 = 1$. נציב ונקבל:

$x^2 - 3x + 2 = \frac{1}{12}sin(1) \sim 0.07$.

נפתור את המשוואה הבאה:

$x^2 - 3x + 1.93 = 0$ ונקבל $x_2 = 0.93$.

כעת נציב באגף ימין את x_2, ונחשב $\frac{1}{12}sin(0.93) = 0.0668$.

נפתור את המשוואה $x^2 - 3x + 2 = 0.0668$.

$x_3 = 0.94$.

באופן דומה, שוב נציב באגף ימין ונמשיך לפתור את המשוואה.

$x^2 - 3x + 2 = \frac{1}{12}sin(0.94)$

$x_4 = 0.94$.

כעת נוכל לעצור. קיבלנו את אותו פיתרון עבור שני האגפים. מאחר ואנחנו רוצים להגיע לדיוק כלשהו, העובדה שיש שוויון בין הפתרונות בשני האגפים תיתן לנו פיתרון די טוב.

האלגוריתם נקרא "אלגוריתם נקודת השבת" והוא עובד באופן הבא:

אם $f(x_n) = x_{n+1}$ היא נקודת שבת של f , כלומר x כך שמתקיים $f(x) = x$.

נרצה לדעת האם בהינתן נקודת השבת x, היא נקודת שבת יציבה, כלומר אם נתחיל קרוב ל-x, האם הסדרה תתכנס ל-x?

מטור טיילור מתקיים $f(x+\epsilon) = f(x) + \epsilon f'(x) + o(|\epsilon|)$

כלומר, $f(x+\epsilon) \approx x + \epsilon f'(x)$

נעביר אגפים ונקבל:

$|f(x+\epsilon) - x| < |\epsilon||f'(x)|$ מי שוויון המשולש ההפוך.

ולכן כאשר $|f'(x)| < 1$, עבור ϵ מספיק קטן נקבל $|f(x+\epsilon) - x| < |\epsilon|$ כלומר, נתקרב ל-x.

$|f'(x)|$ נקרא המכפיל בנקודת השבת של x. נשים לב שמתקיים:

אם $|f'(x)| < 1$ - נקודת השבת יציבה.

אם $|f'(x)| > 1$- נקודת השבת לא יציבה.

אם $|f'(x)| = 1$ נקודת השבת בעלת יציבות גבולית (יכולה להיות יציבה ויכולה להיות לא יציבה).

אם $|f'(x)| = 0$ נקודת השבת מאוד יציבה.

הרעיון הוא לחפש קירוב לפתרון המשוואה $f(x) = 0$. אנחנו מוצאים פונקציה $g(x)$ כך שמתקיים $g(x) = x$ בשורש של f. בשלב ה־n מחשבים את הקירוב $x_{n+1} = g(x_n)$ ־ יש משפטים המבטיחים את ההתכנסות של x_n לשורש . הכי שימושי הוא $|g'(x)| < 1$ בסביבת השורש.

דוגמא מציאת נוסחת איטרציה למציאת פתרון למשוואה $f(x) = x^5 - xe^x + sin(x)$.

נסתכל על המשוואה $f(x) = 0$, כלומר $f(x) = x^5 - xe^x + sin(x) = 0$.

נעביר אגפים, $x^5 - xe^x = -sin(x)$

$x = -\frac{sin(x)}{x^4 - e^x} g(x)$

$f(x) = 0 \iff g(x) = x$

נחשב את הנגזרת של g,

$g'(x) = -\frac{cos(x)\cdot(x^4-e^x)-sin(x)(4x^3-e^x)}{(x^4-e^x)^2}$

עבור x מספיק קטן (קרוב לאפס), $|g'(x)| < 1$. ולכן $g(x_n) = x_{n+1}$ אם נבחר בסביבה מספיק קטנה.

שיטת ניוטון רפסון

בהינתן פונקציה שמחפשים את השורש שלה (נגביל את התחום של הפונקציה למקרה בו יש לה שורש אחד בלבד), אם נבחר נקודה קרובה לשורש, שורש המשיק באותה נקודה יהיה קרוב יותר לשורש שאנחנו מחפשים. לאחר מספר איטרציות נקבל קירוב טוב מספיק. האלגוריתם עובד באופן הבא:

1. נבחר נקודה הקרובה לשורש שאנחנו מחפשים־ לנקודת ההתחלה יש חשיבות.
2. נחשב את שיפוע המשיק בנקודה (הנגזרת של הפונקציה בנקודה).
3. נחשב את משוואת המשיק.
4. נמצא את שורש המשיק (הנקודה בה המשיק חותך את ציר x).

בהינתן $f : [a, b] \to \mathbb{R}$ גזירה בקטע $[a, b]$. נתחיל את האיטריצה בנקודה x_0 . שיפוע המשיק בנקודה זו יהיה $f'(x_0)$. כעת נמצא את משוואת הישר העובר דרך הנקודה $(x_0, f(x_0))$ ושיפועו $f'(x_0)$. המשוואה תהיה $y - f(x_0) = f'(x_0)(x - x_0)$. נמצא את

השורש על־ידי הצבת $y = 0$, ונקבל $x = x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$. נמשיך באופן דומה, כך שהסדרה $\{x_n\}_{n=0}^{\infty}$, עבור $x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$ תתכנס לשורש של f (כאשר נבחר x_0 מתאים).

הרעיון הוא שברזולוציה גבוהה מספיק, הפונקציה דומה לקו ישר (בשביל זה צריך שלפונקציה תהיה נגזרת רציפה).

נסתכל על הפונקציה שראינו קודם לכן, $f(x) = x^2 - 3x + 2 - \frac{1}{12}sin(x)$.

נתחיל מ $x_0 = 1$.

$f'(x) = 2x - 3 - \frac{1}{12}cos(x)$.

$f'(1) = -1.08$

$f(1) = -0.07$

$x_1 = 1 - \frac{f(1)}{f(1)} = 0.93$

וכן הלאה.

דוגמא נרצה למצוא שורש של מספר כלשהו a. נסתכל על הפונקציה $f(x) = x^2 - a$. הפונקציה גזירה על כל הישר, $f'(x) = 2x$. נניח שנרצה למצוא את השורש של 2, אזי $x_0 = 2$, ונמצא את הישר העובר דרך הנקודה $(2, 2)$ ושיפועו 4. $y - 2 = 4x - 8$, $y = 4x - 6$. נציב $y = 0$ ונקבל $x = \frac{6}{4} = 1.5$. הסדרה שלנו תהיה $x_i = x_{i-1} - \frac{f(x_{i-1})}{f'(x_{i-1})} = x_{i-1} - \frac{x_{i-1}^2 - 2}{2x_{i-1}}$

נמשיך באופן דומה, $x_2 = 1.5 - \frac{0.25}{3} = 1.41667$

$x_3 = 1.4167 - \frac{0.006944}{2.832} = 1.41424$

וכן הלאה. נשים לב שכבר באיבר השלישי יש דיוק של 5 ספרות אחרי הנקודה.

משפט אם f גזירה ברציפות ואם x_0 מספיק קרוב לפיתרון האמיתי, השיטה תתכנס לפיתרון.

$Implied\ Valatility$

משוואות בלק שולס:

$$c = \phi(d_1)s - \phi(d_2)ke^{-rt}$$

$$p = \phi(-d_2)ke^{-rt} - \phi(-d_1)s$$

c,p אופציות $call$ ו־put

ϕ־ התפלגות מצטברת של משתנה מקרי נורמלי סטנדרטי.

s מחיר המניה היום.

k הסטרייק (מחיר מימוש).

t פרק הזמן עד הפקיעה.

r ריבית חסרת סיכון.

$d_1 = \frac{ln(\frac{s}{k})+(r+\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}$.

$d_2 = \frac{ln(\frac{s}{k})+(r-\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}$.

σ היא הנדיפות ($volatility$).

נתבונן בנתוני השוק־

c,p ידועים

s,k ידועים

t,r ידועים.

σ לא ידועה אך עם זאת, יש הרבה דרכים להעריך אותה, ונרצה לעשות זאת.

לשם כך נשתמש בשיטה נומרית.

בשלב ראשון־ נניח שהשוק מתמחר את c,p בדיוק לפי נוסחאות בלק־שולס. כלומר, נניח שהמשוואות נכונות והדבר היחיד שמשתנה במשוואה הוא ה־σ. נסתכל על המשוואות ונחלץ מהן את ה־σ המתאימה שהופכת את הנוסחה לנכונה. אם נציב באגף ימין את כל הפרמטרים שאנחנו יודעים פרט ל־σ, אז נוכל לקבל את σ(שהופך את ההנחה שהשוק מתמחר את האופציות לפי בלק שולס לנכונה). ניקח את כל הפרמטרים הידועים מהשוק ונפתור את המשוואה הבאה: $p = \phi(-d_2)ke^{-rt} - \phi(-d_1)s$ כאשר הנעלם היחיד הוא σ.
הפתרון של המשוואה נקרא ה־ Implied volatility והוא יהיה בשיטת ניוטון־רפסון.

הערה בפייתון הפונקציה $newtstep$ מקבלת ניחוש לפיתרון ומחזירה את הניחוש הבא לפי ניוטון רפסון. בהינתן פונקציה f, לה נרצה למצוא שורשים, כלומר, נפתור את המשוואה $f(x) = 0$, נעביר משיק ל־f בנקודה $(x_1, f(x_1))$ ומוצאות את נקודת החיתוך שלו עם ציר x. לאחר מכן נרצה למצוא את נקודת החיתוך של המשיק עם ציר ה־x. נמשיך להריץ את הפונקציה עד תנאי עצירה שנקבע מראש. במקרה שצוין לעיל, הפונקציה היא $f(\sigma)$.

$f(\sigma) = bsput(\sigma) - p$ כאשר $bsput(\sigma)$ הוא מחיר put לפי בלק שולס עם נתוני השוק ו-p הוא מחיר השוק.

על מנת לפתור את הבעיה, נרצה למצוא σ כך ש-$f(\sigma) = 0$.

$\sigma_2 = \sigma_1 - \frac{bsput(\sigma)-p}{bsput'(\sigma)}$.

$bsput'(\sigma)$ הוא למעשה הנגזרת (p הוא קבוע). מכיוון שאנחנו לא יודעות לחשב את הנגזרת, נחשב קירוב לנגזרת.

לפי ההגדרה של הנגזרת, מתקיים $f'(x) = lim_{u\to 0}\frac{f(x+u)-f(x)}{u}$.

כאשר יש גבול כזה, ניקח u קרוב מספיק לאפס ונקבל $f'(x) \sim \frac{f(x+u)-f(x)}{u}$.

בדוגמא ניקח $u = 10^{-5}$ (מניחים שזה מספיק קטן) וממנו נקבל קירוב לנגזרת.

שיעור 2

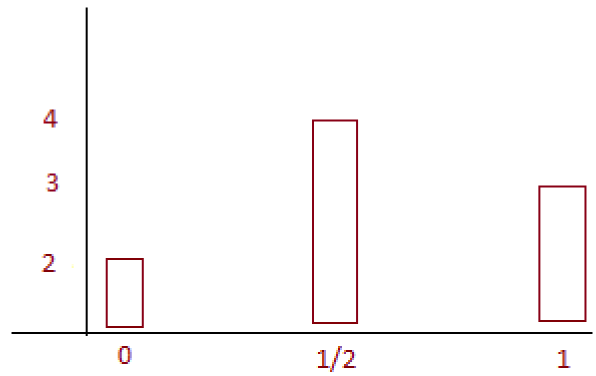
ייצור משתנים מקריים

נרצה לייצר אוסף מספרים בעל תכונות הדומות להתפלגות משתנה מקרי כלשהו.

נייצר אוסף מספרים $x_1, ..., x_n, ...$ בעל תכונות של משתנה מקרי X, כאשר ל-X יש התפלגות כלשהי (למשל $X \sim N(0,1)$). התכונה הבסיסית אליה נצפה מבחינת תכונות האוסף היא שההיסטוגרמה של $x_1, ..., x_n$ תהיה דומה לצפיפות של X.

הערה ההיסטוגרמה היא גרף בו מחלקים את הטווח הנידון לקטעים באורך שווה ומעל כל קטע מציירים עמודה אשר גובהה הוא מספר הערכים מתוך האוסף $x_1, ..., x_n$ הנופלים בתוך הקטע.

נסתכל על הדוגמא הבאה: בהינתן הסדרה $x_1, x_2, x_3, ... = 0, 1, 1, 1, 0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}$, נקבע 3 נקודות ונספר כמה פעמים מופיע כל ערך.



תכונות נוספות שנרצה עבור הסדרה 1. אי תלות בין הדגימות ־ המשמעות של אי תלות היא שאם אנחנו יודעות את הערכים $x_1, x_2, ..., x_n$, אי אפשר להסיק דבר על ערכו של x_{n+1}. תכונה זו לעולם לא תתקבל באופן מלא, אבל המטרה היא לקחת אלגוריתמים שיסתירו את אותה אי תלות. אנחנו נסתפק לרוב בתכונה חלשה יותר:

2. חוסר קורלציה בין האיברים בסדרה.

ייצור מספרים פסאודו־רנדומליים

נקודת המוצא שלנו היא ייצור דגימות של התפלגות אחידה על הקטע $[0,1]$ (התפלגות $U[0,1]$).

תזכורת, אם $X \sim U[0,1]$, אזי,

$$f_X(t) = \left\{ \begin{array}{ll} 1 & t \in [0,1] \\ 0 & Otherwisse \end{array} \right\}$$

הערה $U[0,1]$ היא הבסיס לכל שאר המשתנים המקריים ולמעשה כל האלגוריתמים לייצור דגימות מהתפלגות כלשהי, משתמשים בדגימה של $U[0,1]$.

אלגוריתמים

1. LSG Linear sequential generator. הרעיון של האלגוריתם הוא לייצר סדרה $x_1, ..., x_n, ..$ לפי הכלל הבא־ $x_{n+1} = ax_n + b(mod c)$. כלומר x_{n+1} הוא שארית החלוקה של $ax_n + b$ במספר c כאשר a, b, c הם פרמטרים כלשהם שצריך לבחור מראש. הסדרה המתקבלת היא סדרת מספרים בין 0 ובין $c - 1 + 1$. לאחר חלוקה של כלל המספרים ב־c, נקבל סדרת מספרים בקטע $[0, 1]$.

הפרמטרים a, b, c קובעים למעשה האם הסדרה תהיה "טובה". פירוש הדבר הוא שבין איברי הסדר לא נוכל למצוא קשר כלשהו. הסדרה לא תהיה רנדומלית, אבל בחירה נכונה של הפרמטרים תיתן לנו סדרה פסאודו רנדומלית.

דוגמא נבחר את הפרמטרים הבאים־ $a = 13, b = 0, c = 31$. נשים לב שa, c הם ראשונים (האלגוריתם עובד טוב כאשר הבחירה היא באיברים ראשוניים).

נתחיל עם $x_0 = 1$. ונשתמש בנוסחה $x_{n+1} = 13x_n mod 31$

$x_0 = 1$

$x_1 = 13$

$x_2 = 14$

$x_3 = 27$

נמשיך לחשב עבור $n = 1, 2, ..., 29$.

נקבל סדרה של מספרים בין 0 ובין 30 בלי חזרות ובלי סדר ניכר לעין. האיבר ה־31 בסדרה יהיה שוב 1 ומשם ממשיכים.

נשים לב כי זו הייתה בחירה טובה של a, c כי מילאנו את כל האפשרויות שלנו ־ עברנו על כל הרצף ללא חזרות.

אם נבחר $a = 13.b - 0, c = 32$, נקבל את הסדרה הבאה:

$1, 13, 9, 21, 17, 29, 25, 5, 1, ...$. במקרה זה עברנו רק על 8 מתוך 32 המספרים ולכן הבחירה לא טובה.

LSG מייצר סדרה מחזורית אבל נרצה שיהיה לנו מחזור גדול כך שהקורלציה לא תיראה דרך הסדרה.

2. LSG לא נמצא בשימוש כיום. רוב חבילות התוכנה משתמשות באלגוריתם שנקרא $Mersune\ Twister$ (שנת 2010). גם באלגוריתם זה משתמשים בנוסחה, $x_{n+1} = F(x_n, x_{-1}, ..., x_1, x_0, , A_n, ..., A_0)$. כלומר האיבר האחרון יהיה פונקציה כלשהי של האיברים הקודמים בסדרה ושל משתנים סמויים (לא נגישים למשתמש). בפייתון הגרעין מתעדכן תמיד מאותו מספר (למרות שיש אפשרות לשנות את ה־$seed$).

3. $Random.org$ ־ אלגוריתם בשם $True\ Rand$ המבוסס על קרינה אלקטרומגנטית מהחלל. הרעיון הוא שיש אפשרות למדוד את התנודתיות על־ידי אנטנות רגישות.

מכאן והלאה נניח שיש לנו גישה בלתי מוגבלת לדגימות $U[0,1]$ ונשתמש באלגוריתמים הקיימים בפייתון.

ייצור משתנים מקריים מהתפלגויות שונות

נרצה כעת לייצר דגימות משתנים מקריים מהתפלגויות שונות, למשל מ־ $exp(\lambda)$,$N(\mu,\sigma^2)$ וכן הלאה. לשם כך יש לנו 2 שיטות עיקריות:

1. שיטת הטרנספורמציה/ההיפוך/ההמרה. אלגוריתם זה עדיף עבורנו במידת האפשר (אם כי השימוש בו לא תמיד אפשרי).
2. שיטת הדחייה־ שיטה כללית יותר, פחות יעילה.

שיטת הטרנספורמציה

המטרה שלנו היא לדגום את $X \sim F_X$ בעל התפלגות כלשהי כאשר ברשותנו דגימות של $U[0,1]$, $u_1, ..., u_n, ..$. כאן נרצה לבחור בחוכמה פונקציה $g(t)$ כך שאם $U \sim U[0,1]$, אזי $g(U) \sim F_X$.

במקרה זה, נציב את הדגימות של המשתנה המקרי האחיד, א..$u_1, ..., u_n,$בפונקציה g ונקבל את הדגימות ... $g(u_1), g(u_2),$ של X, מתוך ההנחה ש־$g(U) \sim F_X$. הבעיה העומדת לפנינו היא מציאת הפונקציה g. קיימת דרך אחת טובה־ נסמן $g(t) = F_X^{-1}(t)$ כאשר F_X היא פונקציית ההתפלגות המצטברת של X ו־F_X^{-1} היא ההופכית שלה.

משפט אם $U \sim U[0,1]$, אזי המשתנה המקרי $F_X^{-1}(U)$ מתפלג כמו X.

מתוך המשפט נקבל $g(U) \sim F_X$ כנדרש.

דוגמא $X \sim exp(\lambda)$. $f_X = \lambda e^{-\lambda t}$. $F_X(t) = 1 - e^{-\lambda t}$.

נרצה לדגום את X בשיטת הטרנספורמציה ולשם כך נרצה למצוא את F_X^{-1}.

$e^{-\lambda t} = 1 - F_X(t)$

$-\lambda t = ln(1 - F_X(t))$

$t = -\frac{ln(1-F_X(t))}{\lambda}$.

מצאנו $F_X^{-1} = -\frac{ln(1-s)}{\lambda} = t$, כלומר ביטוי סגור ל-$F_X^{-1}(s)$.

כעת נדגום .. $u_1, u_2,$ מהתפלגות אחידה $U[0,1]$. נציב ב-$F_X^{-1}(s)$ ונקבל דגימות של X.

$x_1 = F_X^{-1}(u_1) = -\frac{ln(1-u_1)}{\lambda}$

$x_2 = F_X^{-1}(u_2) = -\frac{ln(1-u_2)}{\lambda}$

וכן הלאה.

אנחנו נקבל שההתפלגות של .. $x_1, ..., x_n$ זהה להתפלגות האקספוננציאלית.

בעיות בשיטות הטרנספורמציה

1. יש משתנים מקריים X עבורם F_X^{-1} לא קיימת (אפשר לפתור את זה, לא בקורס).
2. הבעיה האמיתית של שיטת הטרנספורמציה היא שיש משתנים מקריים X שעבורם אי אפשר לחלץ את t מהמשוואה $F_X(t) = s$ ולכן אין דרך פשוטה לחשב את F_X^{-1}.

דוגמא למקרה האחרון יהי $X \sim N(0,1)$. אזי $f_X(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$, $F_X(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds$

מאחר ובמקרה זה לא ניתן להתמודד עם האינטגרל, לא ניתן לחלץ את t אם נסמן $\int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds = g$ ולכן אין נוסחה פשוטה ל-$F_X^{-1}(q)$.

למשתנים מקריים נורמליים נמצא שיטות אחרות לפיתרון הבעיה.

הערה לפעמים יש בחירה אחרת על הפונקציה g שלא תהיה בהכרח F_X^{-1}.

למשל:

1. $Box - Mullen$
2. $Marsaglia$

אלגוריתם לייצור $N(0,1)$ על ידי בחירה חכמה של g שהיא לא F^{-1}.

שיטת הדחייה

הרעיון הוא שנרצה לדגום משתנה מקרי X. כאשר קשה לעשות זאת, נוכל לדגום משתנה מקרי Y, הדומה ל־X (כלומר, בעל פונקציית צפיפות קרובה לפונקציית צפיפות של X) וקל יותר לדגימה.

נניח שניתן לדגום את Y לפי שיטת ההיפוך אבל את X אי אפשר. במקרה זה נדגום את Y ובהתאם למה שהתקבל נחליט אם הוא מתאים להיות X. אם לא, נדחה אותו.

אלגוריתם שיטת הדחייה יהיו X,Y משתנים מקריים רציפים עם פונקציות צפיפות f_X, f_Y בהתאמה.

1. מוצאות מספר $c \geq 1$ כך ש־$\frac{f_X(s)}{f_Y(s)} \leq c$ לכל s. בחירת c נעשית על־ידי חישוב אנליטי המבוצע לפני הרצת האלגוריתם. על־פי הערך של c, נחליט האם לקבל או לדחות ערך מסוים (כלומר, נחליט האם אותו ערך מתאים/לא מתאים להיות דגימה של X). את c בוחרים פעם אחת בלבד ואיתו עובדים לאורך כל הרצת האלגוריתם.
2. מייצרות דגימה מY $(y_1, y_2, ...)$.
3. דוגמות Uאחיד, כלומר מייצרות דגימה חדשה ובלתי תלויה של $U[0,1]$, נסמנה ב־u_i.
4. אם מתקיים אי השוויון הבא: $u_i \leq \frac{f_X(y_i)}{cf_Y(y_i)}$ אזי y_i "מתאים" ונציב $u_i = y_i$ והדגימה היא מ־X.
5. אם לא מתקיים אי השוויון, נזרוק את הערך y_i ו־u_i שדגמנו בשלבים 2 ו־3 ונתחיל מחדש בשלב 2.

הערות 1. בסופו של דבר יימצא y_iמתאים ונקבל דגימה מX, x_i. נחזור על שלבים $2-5$ על מנת לייצר סדרת משתנים מקריים $x_1, x_2, ..$.

2. איך נחשב את c? הערך האידיאלי של c יהיה $c = \left[\frac{f_X(s)}{f_Y(s)}\right] sup_{s\in\mathbb{R}}$. כלומר, נרצה למצוא את הערך הקטן ביותר המקיים את אי השוויון מלעיל. מאחר ומדובר על גבול עליון, לא תמיד ניתן לחשב את ערך זה, אבל למעשה, זה גם לא תמיד הכרחי. ככל שנתקרב לאותו ערך, נקבל אלגוריתם יעיל יותר. ניתן לבחור ערך גדול יותר של c בתמורה למחיר של יעילות (במקרה שנבחר c גדול יותר, נדחה יותר ערכים של y_i).

3. כמה פעמים בממוצע נדחה דגימה מ־ Y עד שתימצא אחת מתאימה להיות דגימה מ־X? ההסתברות שדגימה תתקבל היא $p_0 p(U \le \frac{f_X(Y)}{cf_Y(Y)})$. מאחר ויש לנו משתנים מקריים המתאימים ל־ y_i, u_i (בלתי תלויים) נוכל להשתמש בהסתברויות. השאלה עוסקת במספר כשלונות עד הצלחה ראשונה־ כלומר, מספר הפעמים שנדחה y_i כדגימה של x_i עד שנקבל דגימה ההוא משתנה מקרי גאומטרי $Geo(p_0)$ ולכן תוחלתו היא $\frac{1}{p_0}$ שהיא למעשה מספר הפעמים הממוצע שנדחה y_i עד קבלת משתנה מקרי.

לשם חישוב התוחלת, נחשב את p_0:

$$p_0 p(U \le \frac{f_X(Y)}{cf_Y(Y)}) = \int_{-\infty}^{\infty} p(U \le \frac{f_X(Y)}{cf_Y(Y)}/Y = t) f_Y(t)dt =$$

נוסחת ההסתברות השלמה

$$= \int_{-\infty}^{\infty} p(U \le \frac{f_X(t)}{cf_Y(t)}/Y = t) f_Y(t)dt =$$

קיבלנו כאן ביטוי התלוי רק ב־t, כלומר ההבדל היא שכאן מציבים $Y = t$ בתוך הביטוי, ולכן ניתן להוריד את ההתניה הראשונה ולקבל ביטוי התלוי ב־U בלבד:

$$= \int_{-\infty}^{\infty} p(U \le \frac{f_X(t)}{cf_Y(t)}) f_Y(t)dt =_{p(u \le ?)=F_U(?)} = \int_{-\infty}^{\infty} F_U(\frac{f_X(t)}{cf_Y(t)}) f_Y(t)dt =_{F_U(t)=t} =$$

$$= \int_{-\infty}^{\infty} \frac{f_X(t)}{cf_Y(t)} f_Y(t)dt = \int_{-\infty}^{\infty} \frac{f_X(t)}{c}dt$$

$$= \frac{1}{c}\int_{-\infty}^{\infty} f_X(t)dt = \frac{1}{c}$$

אינטגרל על כל הישר של פונקציית צפיפות הוא 1.

קיבלנו $p_0 = \frac{1}{c}$־ כלומר ממוצע מספר הדחיות הוא c. נשים לב שנרצה לדחות כמה שפחות, ולכן ככל ש־c יהיה קטן יותר, ממוצע מספר הדחיות יהיה קטן יותר.

שיעור 3

שיטת הדחייה

נרצה לדגום את X כאשר באפשרותינו לדגום את Y בהינתן קבוע c ככך שמתקיים $\frac{f_x(s)}{f_y(s)} \le c$

.

האלגוריתם 1. דוגמות את Y: $y_1, y_2, \ldots$.

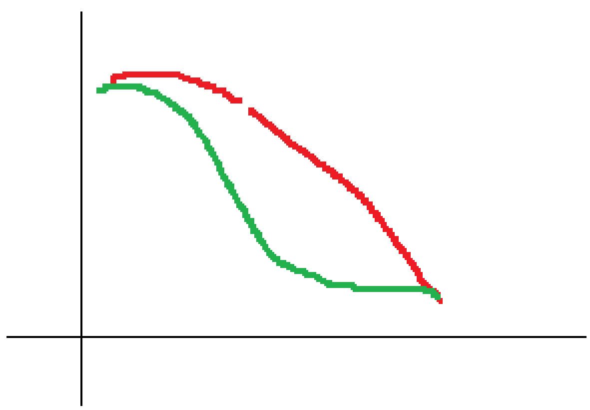
2. דוגמות $U \sim U[0,1]$.

3. בודקות האם $u \le \frac{f_X(y)}{f_Y(y)}$. אם כן, נציב $x = y$, כאשר x היא דגימה של X וסיימנו. אם לא, נחזור להתחלה ונדגום u, y חדשים.

שבוע שעבר ראינו את החשיבות של c. בממוצע נדגום את y c פעמים לפני שנקבל x אחד ראוי , כלומר, נעדיף c קטן.

c מבטא למעשה כמה f_X, f_Y שונות זו מזו. כאשר $c = 1$, $f_X = f_Y$ ולמשתנים המקריים יש אותו התפלגות. המקרה הטוב ביותר עבורנו היה $c = 1$ לפיכך (במקרה זה גם לא נדחה דגימות). מאחר ואנחנו מניחים כי הפונקציות צפיפות שונות זו מזו, $c \neq 1$. איך בכל זאת נוכל לבחור את c? בחירת הערך היא שאלה קשה ולרוב הפיתרונות מתאימים למשתנים מקריים ספציפיים.

דוגמא נתון גרף הפונקציות $h(t) = e^{-\frac{t^2}{2}}$ ו־ $g(t) = 1 - \frac{t^2}{5}$ בקטע $0 < t < 2$.



נציב מספר ערכים לשם קבלת הבנה טובה יותר של הגרף.

עבור $t = 1$ למשל, אמצע הקטע, נקבל $h(1) = e^{-\frac{1}{2}}$, $g(1) = \frac{4}{5}$, קיבלנו $h(1) < g(1)$.

נשים לב כי הגרף האדום מתאים לפונקציה $g(t)$ והירוק לפונקציה $h(t)$.

נרצה לייצר דגימות של משתנה מקרי נורמלי סטנדרטי $N(0,1)$ בקטע $[0,2]$ בעזרת דגימות של מ"מ עם פונקציית הצפיפות $f(t) = \frac{15}{22}(1 - \frac{t^2}{2})$.

אם נשתמש במונחים של שיטת הדחייה, נשים לב ש־X, המשתנה המקרי הקשה יותר לדגימה הוא $N(0,1)$ בהינתן שהוא בקטע $[0,2]$, והמשתנה המקרי Y הוא הקל יותר עם

פונקציית הצפיפות הנתונה $f(t)$.

מתוך הגרף המוצג לעיל, אנחנו רואים שמתקיים אי השוויון $1-\frac{t^2}{5} > e^{-\frac{t^2}{2}}$.

על מנת למצוא את c, נחשב $c = sup_{s\in\mathbb{R}}\left[\frac{f_X(s)}{f_Y(s)}\right]$, $f_Y(t) = \frac{15}{22}(1-\frac{t^2}{5})$.

נכפול בקבוע $\frac{15}{22}$ את אי השוויון לעיל ונקבל: $f_Y(y) = \frac{15}{22}(1-\frac{t^2}{5}) > \frac{15}{22}e^{-\frac{t^2}{2}}$.

נרצה לכפול בקבוע מתאים על מנת שנקבל את $f_X(t)$ באגף שמאל של אי השוויון.

במקרה הנוכחי, $f_X(t) = \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}/[\phi(0)-(1-\phi(2))]$.

הסבר מאחר ונרצה משתנה מקרי נורמלי סטנדרטי בקטע $[0,2]$ נקבל $f_X(t) = \frac{1}{\sqrt{2\pi}(\phi(0)-(1-\phi(2))}e^{\frac{-t^2}{2}}$.

הרעיון הוא שהאינטגרל על כל התחום עליו המשתנה המקרי מוגדר צריך להיות 1,

אנחנו יודעים שמתקיים: $\int_{-\infty}^{\infty} f_X(t)dt = \int_{-\infty}^{\infty}\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt = 1$ (הצבנו $\mu=0,\sigma=1$)

אם נסתכל על האינטגרל על הקטע $[0,2]$ נקבל:

$$\int_0^2 f_X(t)dt = \int_0^2\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt = 1-\int_2^{\infty} f_X(t)dt - \int_{-\infty}^0 f_X(t)dt =$$

$$\int_{-\infty}^0 f_X(t)dt = F_X(0) = \Phi(0) = 1-\Phi(0)$$

$$\int_2^{\infty} f_X(t)dt = 1-\int_{-\infty}^2 f_X(t)dt = 1-F_X(2) = 1-\Phi(2)$$

ולכן, $\int_0^2 f_X(t)dt = \int_0^2\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt = 1-(1-\Phi(2))-(1-\Phi(0)) = \Phi(2)-1+\Phi(0)$

ומכאן נובע כי $N(0,1) \sim \frac{1}{\sqrt{2\pi}(\phi(0)-(1-\phi(2))}e^{\frac{-t^2}{2}} = f_X(t)$ בקטע $[0,2]$.

נכפול בביטוי הזה ונקבל אי שוויון מהצורה

$\frac{15}{22}\frac{1}{\sqrt{2\pi}(\phi(0)-1+\phi(2)}e^{-\frac{t^2}{2}} = \frac{15}{22}f_X(t) < \frac{1}{\sqrt{2\pi}(\phi(0)-1+\phi(2)}f_Y(t)$. ומכאן ניתן למצוא את c (הקבוע המנרמל).

$\frac{f_X(t)}{f_Y(t)} \leq \frac{22}{15\sqrt{2\pi}(\Phi(0)-1+\Phi(2))} = c$.

למה שיטת הדחייה עובדת? אנחנו למעשה דוגמים את המשתנה המקרי $W = Y/U = \frac{f_Y(Y)}{cf_X(Y)}$ כאשר $U \sim U[0,1]$ וכן Y בעל צפיפות $f_Y(t)$. לפי האלגוריתם, Y הופך להיות X בהינתן שאי השוויון מתקיים (U,Y בלתי תלויים).

נרצה כעת להראות מדוע W בעל אותה התפלגות כמו של X.

הוכחה לכך ש־$F_W(t) = F_X(t)$ $F_W(t) = p(W\leq t) = p(Y\leq t/U \leq \frac{f_X(t)}{cf_Y(t)})$

תחת ההנחה ש־ $U \leq \frac{f_X(y)}{cf_Y(y)}$, $W = Y$.

מנוסחת הסתברות מותנית מתקיים : $p(A/B)=p(A\cap B)/p(B)$.

ולכן נקבל:

$$=p(Y\leq t\cap U\leq \frac{f_X(Y)}{cf_Y(Y)})/p(U\leq \frac{f_X(Y)}{cf_Y(Y)})=$$

לפי חישוב מהשיעור הקודם, קיבלנו את שוויון זה:

$$=\frac{1}{\frac{1}{c}}\cdot p(Y\leq t\cap U\leq \frac{f_X(Y)}{cf_Y(Y)})$$

מכיוון שY,U בלתי תלויים, נרצה להתנות באחד מהם ולהיפטר מהם:

$= c\int_{-\infty}^{\infty}p(Y\leq t\cap U\leq \frac{f_X(Y)}{cf_Y(Y)}/Y=s)f_Y(s)ds =$ האחרון לפי נוסחת ההסתברות השלמה.

נציב כעת את s (כפי שעשינו קודם לכן), חילקנו למקרים על־פי Y:

$$=c\int_{-\infty}^{\infty}p(s\leq t\cap U\leq \frac{f_X(s)}{cf_Y(s)}/Y=s)f_Y(s)ds=$$

s,t מספרים ולכן U הוא המשתנה המקרי היחיד בביטוי, ובהתניה יש ביטוי שתלוי רק ב־Y.

$$=c\int_{-\infty}^{t}p(s\leq t\cap U\leq +c\int_{t}^{\infty}p(s\leq t\cap U\leq \frac{f_X(s)}{cf_Y(s)}/Y=s)f_Y(s)ds= \frac{f_X(s)}{cf_Y(s)}/Y=s)f_Y(s)ds$$

כאן חילקנו את הגבולות של האינגרל ל־ $[-\infty,t]$ ול־$[t,\infty]$.

האינטגרל שבו t הוא הגבול התחתון פירושו ש־$t\leq s$, כלומר, המאורע לא מתקיים: $c\int_{t}^{\infty}p(s\leq t\cap U\leq \frac{f_X(s)}{cf_Y(s)}/Y=s)f_Y(s)ds$ ולכן ההסתברות שלו היא אפס (האינטגרל הוא אינטגרל על אפס ולכן מתאפס).

לגביי האינטגרל השני, העובדה שהגבול העליון הוא t אומר ש־$s\leq t$ מתקיים תמיד ולכן ניתן להוריד אותו מתיאור המאורע ־ הוא לא תורם כלום לתיאור המאורע. במקרה זה מה שנקבל הוא שהאינטגרל השני מתאפס והאינטגרל הראשון הפך להיות אינטגרל על כל שאר הדברים,

ונקבל:

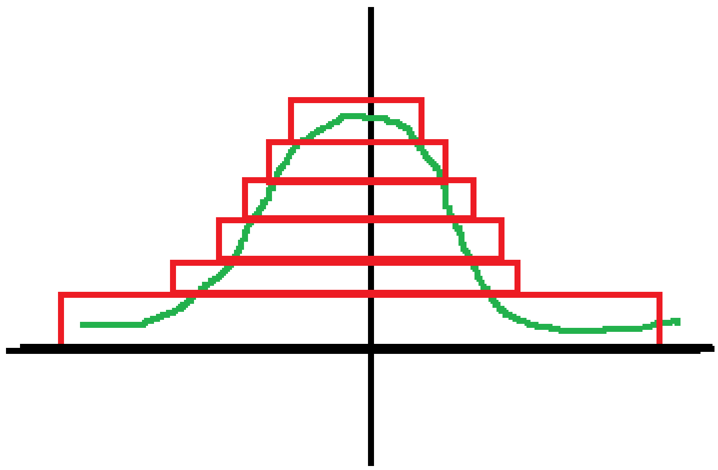
$$=c\int_{-\infty}^{t}p(U\leq \frac{f_X(s)}{cf_Y(s)})f_Y(s)ds=...=c\int_{-\infty}^{t}\frac{f_X(s)}{cf_Y(s)}f_Y(s)ds=\int_{-\infty}^{t}f_X(s)ds= F_X(t)$$

□

(המעבר האמצעי $=...=$ נובע מכך ש־U הוא משתנה מקרי אחיד).

אלגוריתם הזיגורט $(Ziggurat)$

האלגוריתם דומה לשיטת הדחייה. גם כאן נרצה לדגום משתנה מקרי X (משתנה מקרי קשה לדגימה). כאן בונים "זיגורט" מסביב לפונקציית הצפיפות של X.



הערה חשוב שהפונקציה $f_X(t)$ תהיה תמיד בתוך הזיגורט, כלומר, הזיגורט ייבנה מעל לפונקציה.

תכונות נדרשות מהזיגורט 1. יכסה את f_X.

2. שטח כל מדרגה קבוע ושווה לשטח כל מדרגה אחרת.

שתי התכונות הללו הכרחיות, התכונה הבאה רצויה.

3. סך כל שטח הזיגורט יהיה קטן ככל האפשר־ כלומר נרצה להתקרב כמה שיותר לפונקצית הצפיפות.

אלגוריתם הזיגורט למעשה משתמש בשיטת הדחייה על מנת לייצר דגימות של X בעזרת דגימות של משתנה מקרי Y כאשר הצפיפות של Y היא הזיגורט.

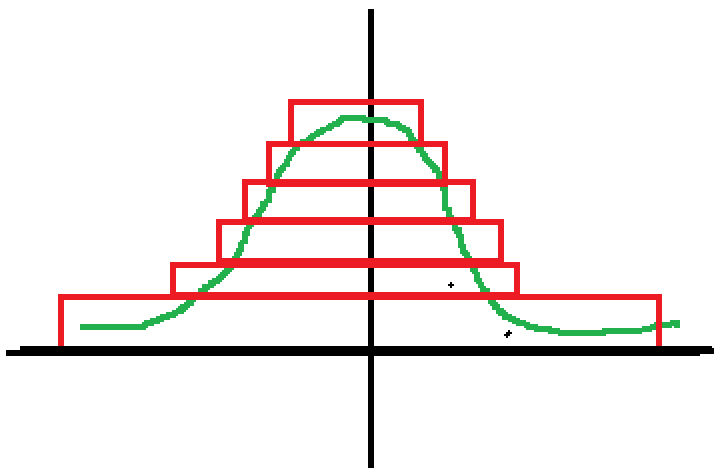
האלגוריתם (בהינתן שהזיגורט כבר בידינו)

1. בוחרות מלבן באקראי. למשל אם לזיגורט יש n קומות, נגריל $1,2,...,n$ אחיד. לאחר מכן, נבחר קומה מתוך הקומות באופן אקראי.

2. אנחנו מייצרות קואורדינטת X אקראית בתוך המלבן הנבחר. למשל אם המלבן הוא: $[a,b]x[c,d]$ נדגום $Z_x \sim U[a,b]$.

3. בודקות האם מעל הקואורדינטה Z_x נפגוש קודם כל את הגג של המלבן, או את הצפיפות של x. כלומר, האם גג המלבן הנבחר יהיה גדול מהפונקציה $f_X(z_x)$.

שתי האפשרויות הללו הן היחידות הקיימות. נסתכל למשל בגרף הבא:

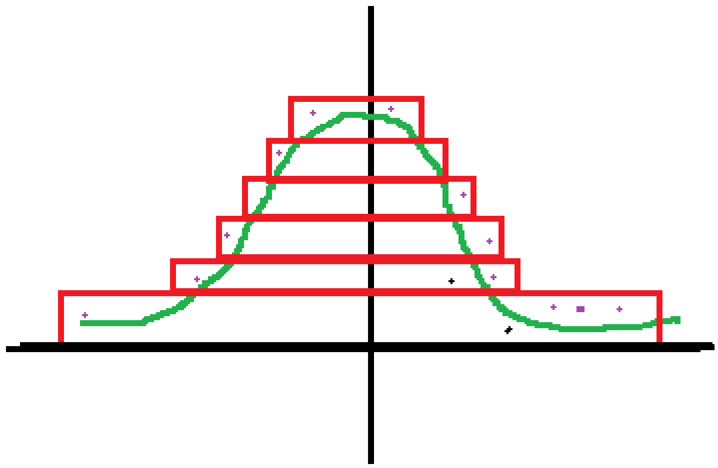


למשל בנקודה השחורה התחתונה, הצפיפות באה לפני הגג, ובנקודה השחורה העליונה, הגג בא לפני הצפיפות. נרצה כמובן לקבל כמה שיותר נקודות בהן הצפיפות תהיה גבוהה מגג הזיגורט.

אם מתקיים כי הגג קטן מ- $f_X(z_x)$, נחזיר את z_x כדגימה של x – כאן אין סיכוי לצאת החוצה, אנחנו בתוך ההתפלגות. אחרת, אם הגג גדול מ-$f_X(z_x)$ נרצה לדעת האם נפלנו בפנים או בחוץ. ולכן כעת נדגום קואורדינטת y: כלומר: $z_y \sim U[c,d]$. אם נפלנו בתוך הצפיפות של X, ניקח את הנקודה, אם $z_y < f_X(z_x)$ כלומר הנקודה (z_x, z_y) בתוך הצפיפות של X, נחזיר את z_x כדגימה. אם $z_y > f_X(z_x)$ נדחה את z_x ונתחיל מחדש ב-1.

נשים לב שאם נבנה את הזיגורט בצורה נכונה, כמעט ולא ניפול מחוצה לו (מקרי הקצה

האלה יתמעטו ככל שנעלה בקומה).



בציור, נדחה את הנקודות הסגולות.

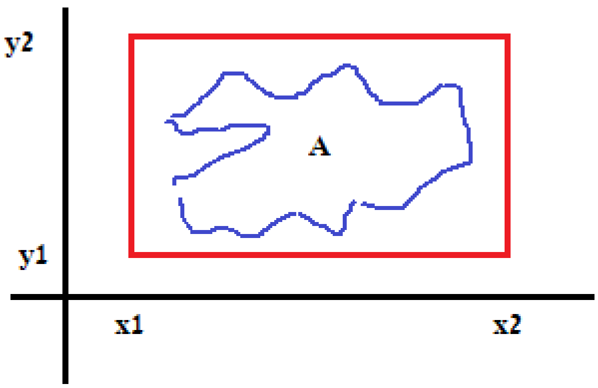
הערה בהמשך נבנה זיגורטים כך שבדיוק כאשר הצפיפות עולה מעל הגג של קומה i, מתחילה הקומה ה־$i+1$. ואז, הבדיקה של 3 היא פשוטה: כל מה שצריך לבדוק הוא האם אני נמצאת משמאל או מימין לקומה הבאה (האם z_x מתחת לקומה הבאה או לא)

שיעור 3 (המשך)

דגימה בתנאי

דגימה של משתנה מקרי אחיד על קבוצה מסובכת נניח שיש לנו קבוצה בעייתית במרחב כאשר לא ניתן לדעת במדויק את הגבולות שלה. נרצה לדגום (X,Y) אחידים בתוך ההקבוצה A. הרעיון הוא להקיף את A במלבן כאשר את גבולות המלבן ניתן למצוא על־ידי (x_i,y_j). קל לדגום (X,Y) אחיד במלבן (באופן דומה למה שעשינו בזיגורט). $X\sim U[x_1,x_2]$ ו־$Y\sim U[y_1,y_2]$. כעת, בהינתן נקודה (X,Y) במלבן, נבדוק עבורה ספציפית אם התנאי שמגדיר את A מתקיים (את זה ניתן לעשות), לרוב מדובר על אי שוויון מסובך. אם התנאי מתקיים נשמור את (x,y) כדגימה של משתנה מקרי אחיד מ־A, אם לא,

נדגום מחדש.



דוגמא (ממבחן) תהיי הקבוצה A הבאה: $A = \{(x,y)|0 \leq x \leq \sqrt{e^{-\frac{x^2}{2y^2}}}\}$.

A מוקפת במלבן $\{(x,y)|0 \leq x \leq 1, -\sqrt{\frac{2}{e}} \leq y \leq \sqrt{\frac{2}{e}}\}$.

נדגום (X,Y) מהמלבן ונציב באי שוויון המגדיר את הקבוצה A.

$Box\ Muller$ באלגוריתם זה למעשה מכניסים משתנים מקריים לעיגול (וכך קל יותר לייצר משתנים מקריים נורמליים).

שיטת מונטה קרלו

נניח כי יש לנו משתנה מקרי מורכב מאוד למשל X מחיר אופציה אמריקאית על אינדקס מניות חודש מעכשיו (כל שוק הטכנולוגיה). המורכבות מגיעה מכך שסכום מניות לא מתנהג כמו מניה בודדת.

נרצה לחשב את התוחלת של X, $E[X]$. למרבה המזל, למרות שקשה לנתח את X, משתנה מקרי בעל התפלגות מורכבת ותלויה בהרבה גורמים, ניתן לייצר דגימות מX.

נסתכל על הדגימות מX, $x_1, x_2, ... x_n$. קיבלנו את הדגימות על־ידי כך ש"גלגלנו" את הזמן קדימה מהיום וקיבלנו שווי לאופציה־ הבאנו את האינדקס למועד הפקיעו של האופציה ו־x_i הוא מחיר המימוש של האופציה בעת המימוש.

נחשב $\bar{X}_n = \frac{x_1+x_2+...+x_n}{n}$. אם נבחר n מספיק גדול אז נקבל ש$\bar{X}_n \approx E[X]$ (לפי חוק המספרים הגדולים $\bar{X}_n \to_{n\to\infty} E[X]$).

כאן לא מספיק לנו לדעת ש$\bar{X}$ קרוב ל$E[X]$ אלא נרצה לדעת גם כמה קרוב.

לפי משפט הגבול המרכזי - עבור n מספיק גדול, $\bar{X} \sim N(E[X], \frac{var(X)}{n})$.

כאשר נדגום ממוצע, נדע כי הוא למעשה בצורת פעמון עם ממוצע $E[X]$ ורוחב פעמון של $\frac{var(X)}{n}$.

כמסקנה מהמשפט, $\bar{X}_n$ יהיה במרחק של $\sqrt{\frac{var(X)}{n}}$ רוב הזמן (68%). כלומר, $\bar{X}_n = E[X] \pm \sqrt{\frac{var(X)}{n}}$.

המספר $\sqrt{\frac{var(X)}{n}}$ מעניין אותי כהערכה לסדר גודל של השגיאה של $\bar{X}_n$ מהתוחלת $E[X]$.

הערה נשים לב כי אם לא נוכל לחשב את התוחלת, לא נוכל לחשב את השונות. מאחר ובתחילת הנושא ראינו כי לא ניתן לחשב את $E[X]$ ולכן, על אחת כמה וכמה, לא ניתן לחשב את $var(X)$.

לפיכך, במקום לחשב את $var(X)$, נשתמש בקירוב (השונות המדגמית), $S^2 = \sum_{i=1}^{n} \frac{(x_i-\bar{x}_n)^2}{n-1}$.

S^2 כמו $\bar{X}$ מחושב מתוך הדגימות.

הערה כאשר עלינו לחשב קירוב ל$E[X]$ יש תמיד לתת הערכה לשגיאה. כלומר, נחשב שני מספרים, $\bar{X}_n, S_n^2$.

דוגמא נרצה לחשב את π.

משפט שטח מעגל הוא רבע π משטח הריבוע החוסם אותו.

הוכחה ניקח מעגל ברדיוס $r = 1$ ונחסום אותו בריבוע. קוטר המעגל הוא $2r = 2$ והוא שווה לאורך צלע הריבוע. שטח הריבוע הוא $2^2 = 4$ ושטח המעגל הוא $\pi r^2 = \pi$ ולכן היחס בין שטח המעגל ובטח הריבוע הוא $\frac{\pi}{4}$.

פיתרון נשתמש במשפט לשם חישוב הערך של π. נסתכל על מעגל ברדיוס $r = 1$ ונחסום אותו בריבוע. נסמן את שטח המעגל ב-A ואת שטח הריבוע ה-B, לפי המשפט, $\frac{A}{B} = \frac{\pi}{4}$. לשם

חישוב π, נרצה לחשב את יחס השטחים בקירוב. נדגום משתנה מקרי אחיד מהריבוע ונסמן 1 אם הוא נמצא בשטח המעגל ו־0 אחרת. כלומר, נדגום $I = \left\{ \begin{array}{cc} 1 & \in A \\ 0 & \in B, \notin A \end{array} \right\}$. ולכן, $E[I]$ הוא יחס השטחים. מאחר ו־I הוא אינדיקטור של מאורע, תוחלת האינדיקטור למאורע היא ההסתברות למאורע, כלומר, $E[I] = p(I = 1)$. $E[I]$ היא ההסתברות שמשתנה מקרי אחיד מתוך B "ייפול" בתוך השטח של A, ומתקיים $E[I] = p(I = 1) = \frac{\int_A dxdy}{\int_B dxdy}$.

על מנת לחשב את π נבצע את השלבים הבאים:

1. נדגום $x_i \sim U[-1,1]$.
2. נדגום $y_i \sim U[-1,1]$.
3. נבדוק האם $x_i^2 + y_i^2 \leq 1$. אם כן, נציב $I_i = 1$, אם לא, נציב $I_i = 0$.
4. נחשב $\bar{I} = \frac{I_1+...+I_n}{n}$.
5. לפי המשפט, $\bar{I} \approx \frac{\pi}{4}$ ולכן הקירוב שלנו לפאי יהיה $\pi = 4\bar{I}$.

בניגוד לשיטת הדחייה, כאן נספור את האפסים.

שיעור 4

אלגוריתמים לייצור מ"מ נורמלי $N(0,1)$

שיטה 1 $N = \sum_{i=1}^{12} u_i - 6$ כאשר $u_i \sim U[0,1]$.

$$E[N] = E[\sum_{i=1}^{12} u_i - 6] = \sum_{i=1}^{12} E[u_i] - 6 = 12 \cdot \frac{1}{2} - 6 = 0$$

$$var(N) = \sum_{i=1}^{12} var(u_i) = 12 \cdot \frac{1^2}{12} = 1$$

קיבלנו משתנה מקרי N שמסכים עם $N(0,1)$ על התוחלת ועל השונות. מאחר וN הוא סכום של 12 משתנים מקריים בלתי תלויים, לפי משפט הגבול המרכזי הוא נורמלי.

שיטת $Box - Muller$ האלגוריתם:

1. יש לדגום $u_1, u_2 \sim U[0,1]$.
2. להציב $\theta = 2\pi u_1$, $R = \sqrt{-2ln(u_2)}$.
3. להציב $z_1 = Rcos(\theta)$, $z_2 = Rsin(\theta)$.

אזי, z_1, z_2 הם בעלי התפלגות $N(0,1)$ ב"ת.

הרעיון הוא שכשעושות טרנספורמציה על מ"מ, הצפיפות משתנה. אם ניקח צפיפות משותפת $f_{U_1,U_2}=1$, ונעשה טרנספורמציה f_{z_1,z_2} נקבל את הצפיפות של משתנה מקרי נורמלי דו מימדי סטנדרטי. הסקיצה היא שנתחיל מנקודות אקראיות בריבוע, והמעבר לR,θ ייתן לנו זווית ורדיוס (מ0 ל2π). לכן, נקבל אחידות מבחינת הפיזור הזוויתי (הזווית θ אחידה) ומבחינת הרדיוס נקבל כי R גדול לא סביר וR קטן סביר (היחס בין הגדול לא סביר וקטן סביר זה בדיוק מה שצריך להתקיים בהתפלגות נורמלית). התפלגות דו מימדית נורמלית היא מעין סדרת מעגלים.

השיפור של $Marsaglia$ באופן דומה לשיטה האחרונה $Box-Muller$ ללא חישוב של $sin(\theta), cos(\theta)$.

נדגום ישירות אחיד על עיגול:

1. דוגמות $u_1,u_2\sim U[0,1]$ בלתי תלויים.
2. מציבות $v_1=2u_1-1, v_2=2u_2-1$. מקבלות $v_1,v_2\sim U[-1,1]$.
3. אם $v_1^2+v_2^2\geq 1$ חוזרות ל1 (שיטת הדחייה).
4. אחרת, אם $v_1^2+v_2^2\leq 1$ עבור $R=v_1^2+v_2^2$ מציבות $z_1=v_1\sqrt{-2ln(R)/R}, z_2=v_2\sqrt{-2ln(R)/R}$ ו־z_1,z_2 הם $N(0,1)$ בלתי תלויים.

מונטה קרלו - שימושים

חישוב אינטגרלים נרצה לחשב את $\int_a^b g(x)dx$. חישוב אינטגרל לא תמיד אפשרי, למשל במקרה בו אין ל־g פונקציה קדומה.

$\int_a^b g(x)dx =(b-a)\int_a^b \frac{1}{b-a}g(x)dx=(b-a)\int_a^b g(x)f_U(x)dx=$

המעבר האחרונה נובע מכך ש־$U\sim U[a,b]$.

קיבלנו אינטגרל של מכפלה של פונקציה בצפיפות, ביטוי המוגדר להיות תוחלת הפונקציה עם הרכבה על המשתנה המקרי

$=E[g(U)](b-a)$.

תזכורת $E[Y^2]=\int t^2 f_Y(t)dt$

$E[ln(Y)]=\int ln(t)f_Y(t)dt$

למעשה, זה נכון לכל g כללית, $E[g(Y)] = \int g(t)f_Y(t)dt$

נשתמש במונטה קרלו עבור אותה תוחלת:

1. נדגום $u_i \sim U[a,b]$.
2. נציב ב־g את u_i.
3. נחשב $\bar{g}(U) = \frac{g(u_1)+...+g(u_n)}{n}$
4. נקבל $\int_a^b g(x)dx \approx \bar{g}(U)\cdot(b-a)$.

הערה עבור המקרה החד מימדי יש שיטות אינטגרציה טובות יותר. כאשר נחשב אינטגרל רב מימדי, שיטת מונטה קרלו היא עדיפה.

אינטגרלים רב מימדיים בהינתן אינטגרל m מימדי, $\int_{a_1}^{b_1}\int_{a_2}^{b_2}...\int_{a_m}^{b_m} g(t_1,t_2,...,t_m)dt_m...dt_2dt_1 =$

$$= [\int\int...\int g(t_1,...,t_m)\frac{1}{(b_2-b_1)...(b_m-a_m)}]\cdot(b_2-b_1)...(b_m-a_m) =$$

$$= (b_2-b_1)...(b_m-a_m)g(t_1,...,t_m)f_{U_1...U_m}(t_1,...,t_m)dt_1...dt_m =$$

$\Pi_{i=1}^m(b_i-a_i)E[g(U_1,...,U_m)]$ כאשר $U_i \sim U[a_i,b_i]$ לכל $i=1,2,...,m$.

את התוחלת נקרב בעזרת ממוצע כפי שעשינו קודם לכן:

נדגום הרבה פעמים ... $u_1 \sim U_1, u_2 \sim U_2,$, נציב ב־$g(u_1,u_2,...,u_m)$ על מנת לקבל מספר יחיד ובסוף נחשב ממוצע $\bar{g}(U_1,...,U_m)$.

נדגום m פעמים משתנה לשם קבלת דגימה בודדת, ונחזור על כך n פעמים.

למעשה יש לדגום מטריצה של משתנים מקריים:

$$\begin{pmatrix} u_{11} & . & . & . & u_{1m} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ u_{n1} & . & . & . & u_{nm} \end{pmatrix}$$

השורה ה־i נותנת לי g_i, כלומר השורה הראשונה תתן לי את g_1, השנייה את g_2 וכן הלאה כאשר האחרונה תתן לי את g_n.

והעמודה ה־j תיתן לי את $g_i(u_j)$.

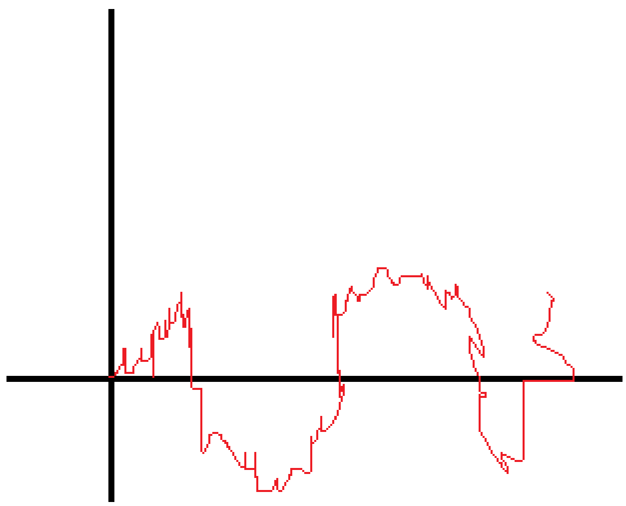
הערה לרוב יהיו יותר שורות מאשר עמודות (נרצה n מספיק גדול).

סימולציה של תהליכים סטוכסטיים

תהליך ווינר (תנועה בראונית) סטנדרטי

הגדרה אוסף משתנים מקריים בעל פרמטר t (W_t) נקרא תהליך וינר סטנדרטי אם מתקיימים התנאים הבאים:

1. $W_0 = 0$.
2. W_t רציף ב־t.
3. הפרשים בלתי תלויים: לכל סדרת זמנים $0 < t_1 < t_2 < ... < t_n$, ההפרשים בין שני זמנים עוקבים הם בלתי תלויים, כלומר $W_{t_n} - W_{t_{n-1}}, ..., W_{t_1} - W_{t_0}$ הם משתנים מקריים בלתי תלויים.
4. ההפרשים נורמליים: לכל $t < s$, $W_s - W_t \sim N(0, s-t)$.



מאחר ואי אפשר לדגום מסלול שלם־ לשם כך צריך ערך ל־W_t בכל נקודה $t > 0$ (אינסוף ערכים), במקום זאת, דוגמות "שלדים" של התהליך.

נתחיל בבחירה של $T > 0$ כלשהו (לרוב מועד הפקיעה של האופציה שאותה נרצה לתמחר). נחלק את ציר הזמן ל־n חלקים שווים (השוויון בגודל החלקים משמעותי להצלחת

האלגוריתם) באורך $h=\frac{T}{n}$. נייצר דגימה של W_t רק בנקודות הנבחרות $(0,h,2h,...,nh=$ $T)$ והתנועה הבראונית תהיה אוסף הדגימות הללו.

דגימת W_t על השלד בהינתן השלד $\{0,h,2h,3h,...,nh=T\}$ נשתמש בתכונות התנועה בראונית.

נתחיל ב־$W_0=0$(תכונה 1). כעת, בהינתן $W_{ih}=c$, נדגום את ההפרש $W_{(i+1)h}-W_{ih}$ לפי תכונה 4, $W_{(i+1)h}-W_{ih}\sim N(0,h)$ ונציב $W_{(i+1)h}=c+\sim N(0,h)$.

משפט דונסקר עבור h קטן מספיק (הרבה צעדים) התהליך הופך לתנועה בראונית (התכנסות תהליכים סטוכסטיים).

בסמסטר הקודם, עם אופציה אירופאית, התחלנו באפס ועשינו רק קפיצה אחת, ייצרנו את W_T בלבד.

כעת נרצה לעבוד עם אופציות תלויות מסלול (אופציות אירופאיות, אופציות אסייתיות) ונרצה את התנועה באמצע המסלול.

באופן מפורט, $w_{1h}=w_0+N(0,h)$

$$w_{2h}=w_{1h}+N(0,h)=w_0+N(0,h)+N(0,h)$$

כאשר שני המשתנים המקריים בלתי תלוים (הפרשים בלתי תלויים).

באופן דומה,

$$w_T=w_{nh}=w_0+N(0,h)+...+N(0,h).$$

עבור n משתנים מקריים בלתי תלויים.

אינטגרלים סטוכסטיים

המטרה שלנו כעת היא לדגום מסלול של אינטגרל סטוכסטי (תהליך איטו).

נתון תהליך לפי המשוואה $dx_t=a(x_t,t)dt+b(x_t,t)dw_t$.

דוגמא $dxt=adt+bw_t$, אזי, x_t היא תנועה בראונית (לא בהכרח סטנדרטית): במקרה זה, a נקרא הסחף $(drift)$, וb הנדיפות $(volatility)$ ושניהם קבועים (נבחר a,b כלשהם).

דוגמא $ds_t = s_t a dt + s_t b dw_t = s_t(ad_t + bdw_t)$. תנועה בראונית גאומטרית, גם כאן a, b קבועים.

דוגמא $dx_t = a(\theta - x_t)dt + bdw_t$ תהליך $vasicek$.

ייצור דגימות מסלולים האפשרות הראשונה, שהיא גם האפשרות הטובה ביותר, היא פיתרון המשוואה הדיפרנציאלית הסטוכסטית. לאחר פיתרון המשוואה, ניתן לייצר מסלול של תנועה בראונית ולהציב בו את הפיתרון.

השיטה בעייתית משום שרוב המשוואות הדיפרנציאליות הסטוכסטיות לא ניתנות לפיתרון. משוואות דיפרנציאליות סטוכסטיות פתירות ניתנות לפיתרון על־ידי הלמה של איטו, עבור משוואות לא פתירות, נשתמש בשיטת הקירוב.

שיעור 5

נרצה לפתור משוואות דיפרנציאליות סטוכסטיות בצורה נומרית.

בהינתן משוואה $dx_t = a(x_t, t)dt + b(x_t, t)dW_t$, נרצה לדגום מסלול של x_t.

תזכורת המשוואה לעיל למעשה מציגה את המשוואה הבאה באופן מקוצר:

$$x_t - x_0 = \int_0^t a(x_{t'}, t')dt' + \int_0^t b(x_{t'}.t')dW_{t'}$$

באופן דומה, ניתן להחליף את הגבול הנתון בs (במקום בתחילת המסלול) ולקבל את המשוואה:

$$x_t - x_s = \int_s^t a(x_{t'}, t')dt' + \int_s^t b(x_{t'}.t')dW_{t'}$$

תחת ההנחה שs מאוד קרוב לt, נסתכל על האינטגרל $\int_s^t a(x_{t'}, t')dt'$.

האינטגרל הוא השטח מתחת ל$a(x_t, t)$ בגבולות s וt. מאחר וקשה לחשב את השטח מתחת לפונקציה $a(x_r, r)$, נחשב את שטח המלבן מתחת לקו הקבוע $a(x_s, s)$ בין s לt. שטח המלבן יהיה $(t-s)a(x_s, s)$ (מכפלת גודל הבסיס של המלבן $t-s$ בגובה המלבן $a(x_s, s)$). בהנחה שכבר דגמנו מסלול של W_t, נטפל בצורה דומה באינטגרל הסטוכסטי, $\int_s^t b(x_s, s)dW_t \approx b(x_s, s)(W_t - W_s)$, כאשר $b(x_s, s)$ מתאר את גובה המלבן, ו־$(W_t - W_s)$ את בסיס המלבן. בסיס במלבן במקרה של האינטגרל הזה הוא משתנה מקרי המתפלג נורמלי, $W_t - W_s \sim N(0, t-s)$.

מכאן כי מתוך המשוואה המקורית $dx_t = a(x_t,t)dt + b(x_t,t)dW_t$, נקבל $x_t \approx x_s + a(x_s,s)(t-s) + b(x_s,s)(W_t - W_s)$. בשיטה זו נשתמש על מנת לדגום מסלול של x_t.

שיטת אוילר־מריומה $(E-M)$

השיטה מייצרת מסלול תהליך איטו בעזרת המשוואה $x_t \approx x_s + a(x_s,s)(t-s) + b(x_s,s)(W_t - W_s)$ מזמן 0 עד זמן T.

האלגוריתם 1. נחלק את הקטע $[0,T]$ לN חלקים שווים באורך h ($h = \frac{T}{N}$).

2. נייצר שלד של מסלול באופן איטרטיבי מ0 ומעלה כאשר בהינתן הערך x_{ih} נחשב את הערך הבא אחריו $x_{(i+1)h}$. נציב $s = ih, t = (i+1)h$ בנוסחה לעיל ונקבל $x_{(i+1)h} = x_{ih} + a(x_{ih},ih) \cdot h + b(x_{ih},ih) \cdot (W_{(i+1)h} - W_{ih})$

נשים לב ש־ $x_{ih}, a(x_{ih},ih), b(x_{ih},ih)$ ידועים וכי $W_{(i+1)h} - W_{ih}$ הוא ההפרש של תנועה בראונית סטנדרטית. לכל ערך i שונה מחשבים הפרש זה לאחרים (אין חפיפה בין הקבוצות). לפיכך, על מנת לייצר את ההפרש, נדגום $N(0,h)$ באופן בלתי תלוי בשאר חישובי האלגוריתם ($W_{(i+1)h} - W_{ih}$ בלתי תלוי ב־$W_{(i+2)h} - W_{(i+1)h}$).

ניצור מערך ערכים אפשריים עבור $x_{(i+1)h}$ ונחשב את הממוצע שלהם כאומדן לתוחלת (כל מערך הוא מקרי ולכן נשתמש במונטה קרלו).

דוגמא נתון נכס בסיס המתנהג $GBM(\mu,\sigma)$, כלומר, $dS_t = S_t(rdt + \sigma dW_t)$, כאשר r היא ריבית חסרת סיכון. נרצה לתמחר אופציה המשלמת את המקסימום של מחיר נכס הבסיס לאורך התקופה עד מועד הפקיעה. נעשה זאת בצורה הבאה:

1. נדגום שלד של מסלול של S_t כפי שראינו קודם, עבור $\{0,h,2h,...,nh = T\}$ נקבל $\{S_0, S_h, ..., S_T\}$.

2. נחשב את המקסימום של S_t בשלד המסלול מהשלב הקודם, $m = max(S_0, S_1, ..., S_T)$.

הדרך הראשונה לחישוב היא דרך אוילר מריומה, כלומר, נשתמש במשוואה שראינו קודם לכן.

הדרך השנייה היא על־ידי שימוש בפיתרון האנליטי של GBM על־מנת לייצר את

המסלול: $S_{(i+1)h} = S_{ih} + rS_{ih} \cdot h + S_{ih} \cdot \sigma N(0,h)$

כאשר rS_{ih}הוא a ־אינטגרל על dt, ו$S_{ih}\sigma N(0,h)$ הוא אינטגרל על dW_t.

3. נחשב את שווי האופציה במועד הפקיעה: $V = (m-k)_+$.

הערה האופציה היא אופציית $call$ אירופאית כאשר נכס הבסיס שלה הוא המקסימום לאורך המסלול.

4. נחזור על שלושת השלבים הקודמים מספיק פעמים.

5. נחשב $\bar{V}$, ממוצע של V על־פני כלל הפעמים בהן חזרנו על השלבים, ואת $\frac{S^2}{\sqrt{n}}$ כקירוב לגודל השגיאה.

הערה ל$\bar{V}$ נצטרך להוסיף את גורם ההיוון e^{-rT}, כלומר, $\bar{V}e^{-rT}$.

הגדרה (השגיאה הדטרמיניסטית) השגיאה הנובעת מכך שאנחנו בוחרות N (מספר קטעים) ומחשבות קירוב ל־x_t, ולא את x_t עצמו, נקראת השגיאה הדטרמיניסטית. אין דרך טובה להעריך אותה מלבד לחזור על החישובים עם ערכי N הולכים וגדלים ולראות שהתשובה לא משתנה.

פתרון אנליטי הדרך השנייה לפיתרון היא על־ידי שימוש בפיתרון האנליטי של GBM, $dS_t = S_t(rdt + \sigma dW_t)$. נפתור באמצעות הלמה של איטו ונקבל, $S_t = S_0 e^{(r-\frac{\sigma^2}{2})t + \sigma dW_t}$.

הפתרון האנליטי עבור הפרש זמנים $[q,t]$ הוא $S_t = S_q e^{(r-\frac{\sigma^2}{2})(t-q) + \sigma(W_t - W_q)}$.

בעזרת הנוסחה האנליטית ניתן לייצר דגימות שלד של מסלול GBM באופן מדויק, כלומר, ללא הטעות דטרמיניסטית הנובעת מקירוב אוילר מריומה לאינטגרלים.

נייצר את המסלול באופן איטרטיבי כאשר $q = ih, t = (i+1)h$ ונניח ש S_{ih} ידוע כאשר מתקיים $h = t - q$.

$S_{(i+1)h} = S_{ih} e^{(r-\frac{\sigma^2}{2})(t-q) + \sigma(W_{(i+1)h} - W_{ih})}$.

נשים לב ש $W_{(i+1)h} - W_{ih}$ הם אותם הפרשים שהשתמשנו בהם באוילר מריומה, כלומר, הם מתפלגים $N(0,1)$ ובלתי תלויים אחד בשני.

הערה כאשר יש פתרון אנלטי למד"ס הוא עדיף־ במקרה זה כאמור לא תהיה לנו טעות דטרמיניסטית. עם זאת, לרוב לא יהיה פתרון אנליטי.

שיעור 5 (המשך)

שיטת אוילר־מריומה

אנחנו מחפשות להעריך את המשתנה המקרי V ־ מחיר אופציה אקזוטית תלוית מסלול. V תלוי במסלול של תהליך איטו x_t כלשהו.

נדגום מסלול של תהליך קרוב ל־x_t ונסמנו ב־y_t. למעשה אנחנו מעריכות את V_y (אופציה עם אותם תנאים על x_t במקום על y_t) ונרצה לדעת מהי השגיאה בהערכה זו.

אנחנו דוגמות $V_y^1, V_y^2, ..., V_y^N$ והאומד (ההערכה) שלנו ל־$E[V_x]$ הוא $\bar{V}_y$.

השגיאה תהיה $|E[V_x] - \bar{V}_y|$־ הערך מוחלט של ההפרש בין המקום בו אני רוצה להיות ובין המקום האמיתי שלי. נרצה כמובן לקבל את $E[V_x]$ אבל מאחר וזה לא אפשרי, נוכל לקבל את $\bar{V}_y$ וההפרש יהיה מעין "שגיאה" בהערכה שלי.

$$error = |E[V_x] - \bar{V}_y| = |E[V_x] - \bar{V}_x + \bar{V}_x - \bar{V}_y|$$

$\bar{V}_x$ הוא ממוצע של M דגימות של V_x (כאשר נכס הבסיס הוא x_t האמיתי).

נשתמש באי שוויון המשולש ונקבל:

$$error = |E[V_x] - \bar{V}_y| = |E[V_x] - \bar{V}_x + \bar{V}_x - \bar{V}_y| \leq |E[V_x] - \bar{V}_x| + |\bar{V}_x - \bar{V}_y|$$

המחובר הראשון בביטוי, $|E[V_x] - \bar{V}_x|$ הוא הטעות הסטטיסטית/הסטוכסטית, שמקורה בכך שמחשבות ממוצע ולא תוחלת. סדר גודל הטעות הוא $o(\frac{1}{\sqrt{M}})$ כאשר M הוא מספר הדגימות. ככל שניקח יותר דגימות, סדר גודל הטעות יהיה קטן יותר.

הביטוי השני $|\bar{V}_x - \bar{V}_y|$ נקרא הטעות הדטרמיניסטית. מקורה של הטעות הדטרמיניסטית הוא בכך שאנחנו דוגמות מסלולים של y במקום מסלולים של x.

גם את סדר הגודל של הטעות הדטרמיניסטית ניתן לדעת והוא $o(\frac{1}{N})$ כאשר N הוא מספר הקטעים המחלקים את המסלול־ מספר נקודות הדגימה בשלד של המסלול (כל קטע כזה מהווה קפיצה במסלול).

הטעות הסטוכסטית היא הטעות הדומיננטית רוב הזמן. מאחר והטעות הדטרמיניסטית פוחתת הרבה יותר מהר מהטעות הסטוכסטית, היא אינה מהווה מקור לבעיות רוב הזמן.

אוילר מריומה במספר מימדים

נסתכל על מערכת המשוואות הדיפרנציאליות הסטוכסטיות הבאה:

$dS_t = S_t(rdt + \sigma_t dW_t)$.

נשים לב שאם $\sigma_t = \sigma$ היה קבוע היינו מקבלים תנועה בראונית גאומטרית.

נניח שמתקיים:

$d\sigma_t = k(\theta - \sigma_t)dt + vdB_t$.

כאשר W_t, B_t הם תהליכי וינר תלויים (נניח שאם B_t עולה, גם W_t עולה).

משמעות התלות בין התהליכים היא שלרוב נתון מקדם מתאם אינפיניטיסימלי של W_t ו־B_t, מספר קבוע ρ כך שהשונות המשותפת של W_t, B_t היא ρ_t (אם נגזור לפי t נקבל את ρ). כלומר, קיים מספר כך ש־ $cov(W_t, B_t) = \rho_t$.

זהו מודל לנכס בסיס המתנהג כ־GBM עם תנודתיות משתנה כאשר התנודתיות מתנהגת לפי מודל $vasicek$.

למערכת המשוואות הבאה:

$dS_t = S_t(rdt + \sigma_t dW_t)$

$d\sigma_t = k(\theta - \sigma_t)dt + vdB_t$

לא קיים פתרון אנליטי ולכן נאלץ להשתמש באחת מהשיטות הנומריות.

נפתור את שתי המשוואות במקביל בעזרת $E - M$. ראשית, נבחר שלד $0 - h - 2h - .... - T = nh$ אנחנו יודעים את S_0, σ_0 ובשניהם נשתמש בו זמנית על מנת לייצר את S_1, σ_1. באופן דומה, נשתמש בערכים S_1, σ_1 על מנת לייצר את הערכים הבאים S_2, σ_2 ונמשיך באופן דומה.

$S_{(i+1)h}$ הוא המחיר בצעד הזמן הבא והוא מקיים:

$S_{(i+1)h} = S_{ih} + (S_{ih}rh + \sigma_{ih}S_{ih}(W_{(i+1)h} - W_{ih})$.

כאשר S_{ih} הוא המחיר בצעד הזמן הקודם, $S_{ih} \cdot r \cdot h$ הוא קירוב ל־$\int S_t r dt$, ו־ $\sigma_{ih} \cdot S_{ih} \cdot (W_{(i+1)h} - W_{ih})$ הוא קירוב ל־$\int S_t \sigma_t dW_t$.

מהמשוואה השנייה נקבל:

$\sigma_{(i+1)h} = \sigma_{ih} + k(\theta - \sigma_{ih})h + v(B_{(i+1)h} - B_{ih})$

ניזכר שההפרשים B_t, W_t תלויים, ונשתמש באותה תלות.

את שתי המשוואות לעיל יש לפתור במקביל ונשתמש בשני הערכים בזמן ה־ih, (S_{ih}, σ_{ih}) על מנת לייצר את הערכים בזמן $i(h+1)$, $(S_{(i+1)h}, \sigma_{(i+1)h})$ כאשר הבנייה נעשית משמאל לימין בקפיצות של h, באופן דומה לבנייה במקרה של משוואה אחת בלבד.

מאחר והתהליכים W_t, B_t תלויים זה בזה, לא נוכל לייצר אותם בעזרת ערכים נורמליים בלתי תלויים כפי שעשינו במקרה של משוואה אחת.

נתונים:

$$Q_1 = (W_{(i+1)h} - W_{ih}) \sim N(0,h)$$

$$Q_2 = (B_{(i+1)h} - B_{ih}) \sim N(0,h)$$

$$cov(Q_1, Q_2) = \rho h.$$

נרצה לייצר משתנים מקריים נורמלים עם שונות משותפת נתונה. נעשה זאת באופן הבא:

נדגום שני משתניים מקריים נורמלים בלתי תלויים, z_1, z_2 כאשר $z_1 \sim N(0,h), z_2 \sim N(0,h)$. נציב $Q_1 = z_1$ ולכן $Q_2 = az_1 + bz_2$ כלומר Q_2 הוא סכום של שני משתנים מקריים נורמלים (הרעש הקודם והרעש החדש) כך שנבחר את a, b באופן זה שיתקיים:

1. $cov(Q_1, Q_2) = \rho h.$
2. $var(Q_2) = h.$

$$cov(Q_1, Q_2) = cov(z_1, az_1 + bz_2) = acov(z_1, z_1) + bcov(z_1, z_2) = ah + b \cdot 0 = ah$$

מכיוון ש־z_1, z_2 בלתי תלויים, השונות המשותפת שלהם היא אפס ולכן הגורם השני מתבטל.

מכאן כי $a = \rho.$

$$var(Q_2) = var(az_1 + bz_2) = var(az_1) + var(bz_2) + 2cov(az_1, bz_2) = a^2h + b^2h + 0 = h(\rho^2 + b^2) = h$$

כלומר, $b^2 = 1 - \rho^2$

$$b = \pm\sqrt{1-\rho^2}.$$

מסקנה נדגום לכל קפיצה בשלד שני משתנים מקריים בלתי תלויים z_1, z_2 ונציב z_1 קפיצה ב־W, $\rho z_1 + \sqrt{1-\rho^2} z_2$ קפיצה ב־B.

שיעור 6

בשיעור הקודם רצינו לפתור את מערכת המשוואות הבאה:

$$dS_t = S_t(rdt + \sigma_t dW_t)$$

$$d\sigma_t = k(\theta - \sigma_t)dt + vdB_t$$

כאשר $\rho_t = cov(W_t, B_t)$.

כתבנו את משוואות אוילר־מריומה והגענו למסקנה שאת ההפרשים: $W_{(i+1)h} - W_{ih}$ ו־$B_{(i+1)h} - B_{ih}$ יש לייצר יחד על סמך שני משתנים מקריים נורמלים z_1, z_2.

מבחינה מעשית, בקובץ $multiD$, מתמחרות אופציית $Call$ על ממוצע של 2 מניות.

S_1, S_2 מתנהגות כ־GBM:

$$dS_1 = rdt + \sigma_1 dW_t^{(1)}$$

$$dS_2 = rdt + \sigma_2 dW_t^{(2)}$$

מקדם המתאם של $W_t^{(1)}, W_t^{(2)}$ הוא ρ.

נרצה לדעת מהו מחיר האופציה ביחס לערך ρ.

שתי המשוואות בעלות אותה ריבית (נניח כי קיימת מידה חסרת סיכון).

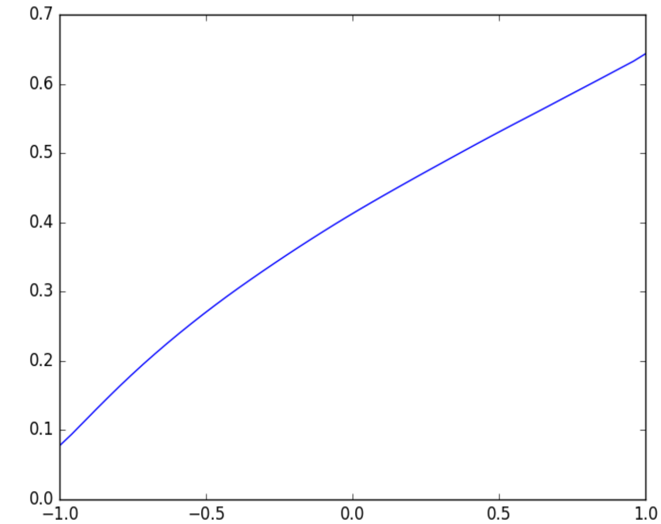
נתונים מספר פרמטרים, ריבית של 5% בשנה ותנודתיות. נריץ מקדם מתאם בין התנועות הבראוניות השונות (ניקח ערכים במרחקים שווים בקטע $[-1, 1]$. נרצה לראות את מחיר האופציה כפונקציה של מקדם המתאם.

ראינו בשיעור הקודם את הנוסחאות הבאות:

$$W_t^{(1)} = z_1, W_t^{(2)} = \rho z_1 + \sqrt{1-\rho^2} z_2.$$

קיבלנו את הגרף הבא המתאר את מחיר האופציה כתלות בקורלציה בין התהליכים:

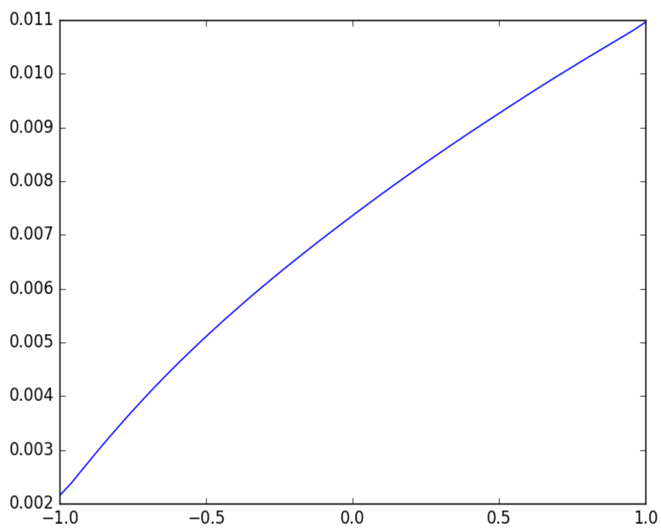
Option price as function of correlation



התוצאה אינטואיטיבית שכן קורלציה גבוהה יותר פירושה סיכון גבוה יותר ולכן מחיר גבוה יותר.

באופן דומה, נסתכל על השגיאה במחיר האופציה כתלות בקורלציה בין התהליכים:

Option price error as function of correlation



אם הנכס מסוכן יותר, יש יותר אפשרויות ולכן המחיר יעלה.

מה עושים עם 3 תנועות בראוניות תלויות? עד כה היינו במערכת של 2 רעשים וכעת נרצה להוסיף אחד נוסף.

נסתכל על המשוואות הבאות:

$$dS_1 = ...sdW$$

$$dS_2 = ...dW^{(2)}$$

$$dS_3 = ...dW^{(3)}$$

כאשר $W_t^{(1)}, W_t^{(2)}, W_t^{(3)}$ תלויים.

במקום מקדם המתאם, ניקח את מטריצת מקדמי המתאם (האינפיניטיסימלי):

$$\rho = \begin{pmatrix} 1 & \rho_{21} & \rho_{31} \\ \rho_{12} & 1 & \rho_{32} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}$$

כאשר:

$$\rho_{ij} \cdot t = cov(W_t^{(i)}, W_t^{(j)})$$

נרצה לייצר הפרשים של שלושת התנועות במקביל:

$$W_{(i+1)h}^{(1)} - W_{ih}^{(1)} = z_1$$

$$W_{(i+1)h}^{(2)} - W_{ih}^{(2)} = az_1 + bz_2$$

$$W_{(i+1)h}^{(3)} - W_{ih}^{(3)} = cz_1 + dz_2 + ez_3$$

כלומר, כל הפרש הוא קומבינציה לינארית של משתנים מקריים נורמלים ($N(0,h)$).

הערה בפייתון, הפקודה $numpy.linalg.cholesky$ מקבלת את מטריצת השונויות המשותפות ומחזירה מטריצה עם המקדמים המתאימים:

$$.cholesky(cov) = \begin{pmatrix} 1 & 0 & 0 \\ a & b & 0 \\ c & d & e \end{pmatrix}$$

לאחר מכן נכפול ונקבל:

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ a & b & 0 \\ c & d & e \end{pmatrix} = \begin{pmatrix} dW_t^{(1)} \\ dW_t^{(2)} \\ dW_t^{(3)} \end{pmatrix}$$

כאשר $z_1.z_2, z_3$ משתנים מקריים בלתי

תלויים, $z_i \sim N(0,h)$ ו $\begin{pmatrix} dW_t^{(1)} \\ dW_t^{(2)} \\ dW_t^{(3)} \end{pmatrix}$ הוא הרעש התלוי שנציב בנוסחאות אוילר־מריומה.

הערה השיטה עובדת באופן זהה לגביי מספר רעשים גדול יותר, כלומר מכניסים מטריצת cov, מקבלים חזרה מטריצה $A = cholesky(cov)$,

$$\begin{pmatrix} dW_t^{(1)} \\ ... \\ dW_t^{(n)} \end{pmatrix} = A \cdot \begin{pmatrix} z_1 \\ ... \\ z_n \end{pmatrix}$$

A היא המטריצה ה"מערבלת" את הרעשים והיא למעשה מכילה את המקדמים הנכונים המשלבים נכון את המשתנים המקריים z_i כעל מנת ליצור את הcov הדרושים.

הלמה של איטו

בהינתן המשוואה $dx_t = a(t,x)dt + b(t,x)dW_t$ וידוע ש־$y_t = g(t,x_t)$, כלומר, y_t מוגדר בעזרת x. לפי הלמה של איטו, y מקיים את המד"ס:

$dy_t = (g_t + ag_x + \frac{1}{2}b^2 g_{xx})dt + bg_x dW_t$ כאשר $g_t = \frac{\partial g}{\partial t}$, $g_x = \frac{\partial g}{\partial x}$, $g_{xx} = \frac{\partial^3 g}{\partial x^2}$.

דוגמא $a = 0, b = 1$ ולכן $dx_t = dW_t$, אזי $x_t = W_t$.

נגדיר $y_t = x_t^2 = W_t^2$ ולכן $g(t,x_t) = x_t^2$.

$g_t = 0$ (t לא מופיע בנוסחה), נחשב $g_x = 2x, g_{xx} = 2$.

נציב בלמה ונקבל:

$dy_t = dt + 2xdW_t$.

נרצה לחשב את האינטגרל של $2xdW_t$.

אזי, נכתוב את הפיתרון בצורה האינטגרלית שלו:

$y_t - y_0 = \int_0^t ds + \int_0^t 2W_s dW_s$.

ידוע ש־$y_t = W_t^2$, $y_0 = W_0^2 = 0$.

נציב ונקבל:

$$W_t^2 - W_0^2 = t + 2\int_0^t W_s dW_s$$

$$\int_0^t W_S dW_s = \frac{W_t^2 - t}{2} = W_t dW_t$$

פיתרון מד"ס על־ידי הלמה של איטו הוא מהסוף להתחלה. כלומר, בהינתן y_t פיתרון של מד"ס נתונה, נציב אותו ונקבל.

ההגדרה של $\int_0^t W_s dW_s$ היא $W_t dW_t$.

מצאנו ש-$\frac{W_t^2 - t}{2}$ הוא פתרון אנליטי ל-$W_t dW_t$.

דוגמא $a = rx_t, b = \sigma x_t$

$dx_t = x_t(rdt + \sigma dW_t)$ כלומר, x_t הוא GBM.

נסתכל כעת על $y_t = lnx_t$

$g(x_t) = ln(x_t)$

$g_t = 0, g_x = \frac{1}{x_t}, g_{xx} = -\frac{1}{x_t^2}$

לפי הלמה של איטו נקבל:

$dy_t = (g_t + ag_x + \frac{1}{2}b^2 g_{xx})dt + bg_x dW_t$

$dy_t = (rx_t \cdot \frac{1}{x_t} + \frac{1}{2}\sigma^2 x_t^2 \frac{-1}{x_t^2})dt + \sigma x_t \frac{1}{x_t} dW_t = (r - \frac{1}{2}\sigma_t^2)dt + \sigma dW_t$.

זו מד"ס של תנועה בראונית לא סטנדרטית:

$y_t - y_0 = \int_0^t (r - \frac{1}{2}\sigma^2)ds + \int_0^t \sigma dW_s = (r - \frac{1}{2}\sigma^2)t + \sigma \int_0^t dW_s = W_t$.

מסקנה $y_t = y_0 + (r - \frac{1}{2}\sigma^2)t + \sigma W_t$

נציב $y_t = ln(x_t)$, נפעיל אקספוננט על שני הצדדים. נקבל:

$x_t = exp(y_0 + (1 - \frac{1}{2}\sigma^2)t + \sigma W_t) = exp(y_0) \cdot exp((1 - \frac{1}{2}\sigma^2)t + \sigma W_t) = x_0 e^{((1 - \frac{1}{2}\sigma^2) + \sigma W_t)}$

זהו הפיתרון האנליטי של התנועה הבראונית הגאומטרית.

דוגמא $dx_t = \lambda(x - x_t)dt + \sigma dW_t$, x_t הוא תהליך $vasicek$.

נסמן $y_t = e^{\lambda t} x_t = g(t, x_t)$

$g_t = \lambda e^{\lambda t} x_t = \lambda y_t$

$g_x = e^{\lambda t}, g_{xx} = 0$

לפי הלמה של איטו, $dy_t = (\lambda y_t + \lambda(c - x_t)e^{\lambda t} + 0)dt + \sigma dW_t$

$dy_t = \lambda c e^{\lambda t} dt + \sigma e^{\lambda t} dW_t$

$y_t - y_0 = \int_0^t \lambda c e^{\lambda s} ds + \int_0^t \sigma e^{\lambda s} dW_s$

$$e^{\lambda t}x_t - e^{\lambda 0}x_0 = c(e^{\lambda s}|_0^t) + \int_0^t \sigma e^{\lambda s} dW_s$$

$$x_t = e^{-\lambda t}x_0 + c(1-e^{-\lambda t}) + e^{-\lambda t}\int_0^t \sigma e^{\lambda s} dW_s$$

$$x_t = e^{-\lambda t}(x_0 - c) + c + \int_0^t \sigma e^{\lambda(s-t)} dW_s$$

זהו פתרון חצי אנליטי של וסיצ'ק (עדיין מכיל אינטגרל סטוכסטי).

כאשר נשתמש במודל וסיצ'ק יש להשתמש בשיטת אוילר־מריומה רק עבור החלק הסטוכסטי $\int_0^t \sigma e^{\lambda(s-t)} dW_s$ ולהציב את הערכים המתקבלים ב־x_t.

$$E[x_t] = E[e^{-\lambda t}(x_0 - c) + c] + = E[\int ..dW_x] = e^{-\lambda t}(x_0 - c) + c$$

המחובר הראשון הוא ההשפעה של נקודת המוצא x_0 ונשים לב כי ההשפעה דועכת מהר. ו־c הוא הממוצע שוסיצ'ק חוזר אליו.

הערה תוחלת של אינטגרל סטוכסטי היא תמיד אפס ולכן החלק הסטוכסטי מתבטל.

שיטת אוילר מריומה

כיצד נוכל להקטין את הטעות הדטרמיניסטית? כיצד נוכל לייצר מסלולים של תהליך דומה יותר ל־x_t?

נקודת המוצא־ אנחנו רוצות לייצר מסלולים של x_t הנתון על־ידי $dx_t = a(x_t)dt + b(x_t)dW_t$.

נשתמש בלמה של איטו עבור a, b, $y_t = a(x_t), z_t = b(x_t)$.

a, b הן פונקציות של משתנה אחד וללא t ולכן הנגזרת שלהן לפי t היא אפס.

$$da(x_t) = (0 + a' \cdot a + \tfrac{1}{2}a'' \cdot b^2)dt + ba'dW_t$$

$$db(x_t) = (b'a + \tfrac{1}{2}b''b^2)dt + bb'dW_t.$$

נעביר אגף בשתי המשוואות:

$$da(x_t) = a(x_t) - a(x_0)$$

$$db(x_t) = b(x_t) - b(x_0)$$

$$a(x_t) = a(x_0) + (a'(x_t)a(x_t) + \tfrac{1}{2}a''(x_t)b^2(x_t))dt + b(x_t)a'(x_t)dW_t$$

באופן דומה עבור $b(x_t)$:

$$b(x_t) = b(x_0) + (b'(x_t)a(x_t) + \tfrac{1}{2}b''(x_t)b^2(x_t))dt + b(x_t)b'(x_t)dW_t$$

נציב את שתי המשוואות הללו במשוואה של x_s ונקבל:

$dx_s = (a(x_0) + (a'a + \frac{1}{2}a''b^2)dt + ba'dW_t)ds + (b(x_0) + (b'a + \frac{1}{2}b''b^2)dt + bb'dW_t)dW_s$

$.x_s = a(x_0)ds + b(x_0)dW_s + (bb'dW_tdW_s) + ...dW_tds + ...dtdW_s + ...dtds$

קיבלנו טור טיילור סטוכסטי, כאשר המשתנים האחרונים במשוואה ניתנים להזנחה.

אם נחשב את $dx_s = x_{(i+1)h} - x_{ih}$ (ההפרש בין הזמנים הוא h), נוכל לומר:

$dtds \approx h^2$ ומאחר ונבחר את h להיות קטן, h^2 יהיה קטן אף יותר.

$dW_t \approx W_{(i+1)h} - W_{ih} \sim N(0,h)$ ולכן $dW_t \approx \pm\sqrt{h}$.

זה אומר ש־$dtdW_t = h\sqrt{h} = h^{\frac{3}{2}}$ עדיין כאן.

אבל $dW_sdW_t \approx (\sqrt{h})^2 = h$, כלומר, ניקח בחשבון את הגורם $bb'dW_tdW_s$.

נשמיט מהמשוואה האחרונה את כל הרכיבים אחרי $bb'dW_tdW_s$ ונפעיל עליה את שיטת EM:

$.dx_t = a(x_0)ds + b(x_0)dW_s + b(x_t)b'(x_t)dW_tdW_s$

$x_s - x_t = \int_0^s a(x_0)dt + \int_0^s b(x_0)dW_r + \int_0^s[\int_0^r b(x_t)b'(x_t)dW_t]dW_r$

יש לנו אינטגרל סטוכסטי כפול, אך אם נסתכל על האינטגרל הראשון והשני, למעשה קיבלנו אותם במתנה ונוכל לכתוב:

$.x_s - x_t = a(x_0)s + b(x_0)(W_s - W_0) + \int_0^s[\int_0^r b(x_t)b'(x_t)dW_t]dW_r$

באופן דומה לEM, האינטגרל הסטוכסטי הכפול הוא להיות קבוע עם הנחה.

$\int_0^s[\int_0^r b(x_t)b'(x_t)dW_t]dW_r \approx \int_0^s\int_0^r b(x_0)b'(x_0)dW_tdW_r$

ולכן נקבל:

$x_s - x_t \approx a(x_0)s + b(x_0)(W_s - W_0) + \int_0^s\int_0^r b(x_0)b'(x_0)dW_tdW_r$ כאשר אנחנו מניחות ש־s מספיק קטן.

התוצאה הסופית תהיה:

$x_s - x_0 = a(x_0)s + b(x_0)(W_s - W_0) + b'(x_0)b(x_0) \cdot \int_0^s\int_0^r dW_tdW_r$

באשר לביטוי האחרון, ראינו שמתקיים:

$\int_0^s[\int_0^r dW_t]dW_r = \int_0^s W_rdW_r = \frac{1}{2}(W_s^2 - s)$

ולכן נקבל:

$x_s - x_0 \approx a(x_0)s + b(x_0)(W_s - W_0) + b'(x_0)b(x_0) \cdot (\frac{1}{2}(W_s^2 - s))$ ־ נוסחת המעבר

מזמן $t=0$ לזמן $t=h$.

במשוואה זו נשתמש בתור הבסיס לשיטה הנומרית הבאה שלנו, שיטה נומרית מדויקת יותר מ־EM הנקראת שיטת מילשטיין.

שיעור 7

וסיצ'ק $dx_t=a(c-x_t)dt+bdW$

$x_t=c+(x_0-c)e^{-at}+\int be^{(s-t)a}dW_s$

כפי שראינו, החלק הראשון הוא החלק הדטרמיניסטי והוא ניתן לפיתרון. לעומת זאת, החלק השני הוא סטוכסטי, לא ניתן לפתרון.

נסמן את החלק השני בy_t ונבצע EM.

$y_t=\int be^{(s-t)a}dW_s$

$x_t=c+(x_0-c)e^{-at}+\int be^{(s-t)a}dW_s$

EM לy_t : $dy_s=be^{as}dW_s$.

$y_{(i+1)h}=y_{ih}+be^{aih}(W_{(i+1)h}-W_{ih})$

נעשה צעד של EM ונכניס אותו בפיתרון האנליטי.

שיטת מילשטיין

$dx_t=a(x_t)dt+b(x_t)dW_t$.

אזי, $x_t-x_0\approx a(x_0)t+b(x_0)W_t+b(x_0)b'(x_0)\frac{1}{2}(W_t^2-t)$.

ממשוואה זו נקבל את שיטת מילשטיין. כמו EM רק שמשוואת המעבר מih ל$(i+1)h$ משתנה.

האלגוריתם 1. נבנה שלד שעליו נדגום מסלול של x_t. בעזרת x_0 נקבל x_h ובעזרתו נקבל x_{2h} וכן הלאה כך שלמעשה נקבל את פרק הזמן מ0 ועד T עבור N קפיצות בגודל h.

נדגום את x_t בנקודות השלד בלבד, כלומר בנקודות $(h,2h,3h,...)$.

2. בהינתן x_{ih}, נקבל $x_{(i+1)h}=x_{ih}+a(x_{ih})h+b(x_{ih})(W_{(i+1)h}-W_{ih})+$

עד כה קיבלנו את EM, וכאן נכנסת תוספת מילשטיין:

$+b(x_{ih})b'(x_{ih})\cdot\frac{1}{2}(W_{(i+1)h}-W_{ih})^2-h$.

היתרון המרכזי הוא שהמשוואה מדויקת יותר מEM (סדר גודל שגיאה $o(\frac{1}{N^2})$).

החיסרון הוא שיש צורך לחשב את b'. לפעמים לא ניתן לעשות זאת בצורה אנליטית ואז החישוב הנומרי גוזל זמן (לפעמים גם חישוב הנגזרת האנליטית גוזל זמן רב).

רוב הזמן (משום שהשגיאה הסטוכסטית היא מסדר גודל של $\frac{1}{\sqrt{m}}$ כאשר m הוא מספר הסימולציות. m זהה בין 2 השיטות, הקצב איטי), השגיאה הדטרמיניסטית איננה הגורם המגביל ולכן נעדיף להשתמש בEM.

הורדת השגיאה הסטוכסטית אנחנו דוגמות משתנה מקרי X מסובך, למשל מחיר אופציה על אינדקס, $x_1, x_2, ...x_n$ ונעריך את התוחלת $E[X]$ על־ידי $E[X] = \bar{x} \pm \frac{\sqrt{S^2}}{\sqrt{m}}$.

S^2 היא השונות המדגמית ושווה ל$S^2 = \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2$.

במקום השגיאה הנ"ל נחשב $E[X] = \tilde{x} \pm \frac{\tilde{\sigma}}{\sqrt{m}}$ כאשר $\tilde{x}$ הוא אומד אחר ל$E[X]$, ו$\tilde{\sigma}$ היא סטיית התקן שלו.

$\tilde{x}$ הוא אומד מתוחכם יותר מהממוצע $\bar{x}$ וסטיית התקן שלו צריכה להיות קטנה יותר מ-$\sqrt{S^2}$ כך שנוכל לקבל הערכה טובה יותר ל$E[X]$.

השיטות

משתנים אנטיתטיים (*antithetic variables*) אם נדגום m משתנים מקריים $N(0,1)$, למשל על מנת לייצר מסלולים של W_t, ונוסיף עוד M סימולציות חדשות עם המספרים הנגדיים לדגימות $x_1, x_2, ...$. כך שנקבל את הדגימות $.... - x_2, -x_1, x_1, x_2, ...$

רעיון השיטה הוא הוספת M סימולציות חדשות.

יתרון השיטה: מרוויחים עוד M סימולציות ללא ייצור של משתנים מקריים חדשים.

חסרונות השיטה:

1. $2M$ הדגימות הן תלויות, מה שעלול להשפיע על הדיוק.
2. לרוב צוואר הבקבוק הוא לא דגימת המשתנים המקריים, ולכן השיטה לא ממש אפקטיבית.

משתנים משותפים (*Common Rvs*) עיקרון השיטה הוא שאם אנו מבצעים מספר סימולציות דומות (אותה סימולציה עבור ערכים שונים של פרמטר כלשהו), רצוי להשתמש באותן דגימות

של המשתנה המקרי.

דוגמא נרצה להעריך את ה־Δ של מניה כלשהי. ניזכר בהגדרה, $\Delta = \frac{dv}{ds_0}$ כאשר v הוא מחיר האופציה וs_0 הוא המחיר ההתחלתי של נכס הבסיס, כלומר Δ היא הנגזרת של שווי האופציה ביחס למחיר ההתחלתי של נכס הבסיס.

$\frac{dv}{ds_0} = lim_{h\to 0}\frac{v(s_0+h)-v(s_0)}{h}$.

נניח כעת ש־h מספיק קטן ונסמן $\frac{dv}{ds_0} \approx \frac{v(s_0+h)-v(s_0)}{h}$.

על מנת להעריך את Δ נדגום את הביטוי $\frac{v(s_0+h)-v(s_0)}{h} \approx \frac{dv}{ds_0}$ מספיק פעמים (נסמן כל דגימה ב־$1 \le i \le M$) ונחשב את ממוצע הדגימות:

$\bar{\Delta} = \frac{1}{M}\sum_{i=1}^{M}\Delta_i = \frac{1}{M}\sum_{i=1}^{M}\frac{1}{h}(v_i(s_0+h)-v_i(s_0))$.

על מנת לחשב את $v_i(s_0)$ ואת $v_i(s_0+h)$ נדגום מסלול i של נכס הבסיס המתחיל בs_0 ומסלול i של נכס הבסיס המתחיל ב־s_0+h. נרצה לייצר את המסלולים כמה שיותר דומים זה לזה, כלומר בעזרת אותו רעש רקע (אותן דגימות עבור W_t).

בחזרה ל $\bar{\Delta}$:

$$\bar{\Delta} = \frac{1}{M}\sum_{i=1}^{M}\Delta_i = \frac{1}{M}\sum_{i=1}^{M}\frac{1}{h}(v_i(s_0+h)-v_i(s_0)) = \frac{1}{h}[\frac{1}{M}\sum_{i=1}^{M}\frac{1}{h}(v_i(s_0+h)-\frac{1}{M}\sum_{i=1}^{M}v_i(s_0)]) =$$

$$= \frac{1}{h}(\overline{v(s_0+h)}-\overline{v(s_0)}) \approx \frac{1}{h}(E[v(s_0+h)]+\frac{dZ}{\sqrt{m}}-E[v(s_0)]+\frac{dZ_2}{\sqrt{m}}) = E[\Delta]+\frac{d}{\sqrt{m}}(Z_1-Z_2)$$

$\frac{dZ_1}{\sqrt{m}}, \frac{dZ_2}{\sqrt{m}}$ הן השגיאות בחישוב ו$\frac{d}{\sqrt{m}}(Z_1-Z_2)$ היא השגיאה של $\bar{\Delta}$.

משתנה בקרה ($Control\ variable$) נרצה להעריך את המסגרת של משתנה מקרי X כאשר ידוע כי הוא מתואם עם משתנה מקרי Y עבור Y פשוט יותר עם תוחלת ידועה.

לרוב נתייחס ל־X כעל מחיר האופציה על נכס הבסיס עם משוואה דיפרנציאלית סטוכסטית בעייתית ול־Y כעל מחיר אותה אופציה על נכס הבסיס GBM, $E[Y]$ ידועה לפי בלק־שולס.

השיטה נדגום את X ואת Y בעזרת משתנים מקריים משותפים, נשווה את $\bar{Y}$ ל־$E[Y]$ ובהתאם אליו נתקן את X כך שהקירוב שנקבל לתוחלת של X יהיה $\bar{X}$ פלוס תיקון קטן התלוי ב־$\bar{Y}$.

האלגוריתם 1.נדגום $x_1, x_2, ..., x_n$ דגימות מX ו־$y_1, y_2, ..., y_n$ דגימות מY, יש לנו $2n$ דגימות.

2.נחשב את הממוצעים, $\bar{X} = \frac{\sum_{i=1}^{n} x_I}{n}, \bar{Y} = \frac{\sum_{i=1}^{n} y_I}{n}$. נסמן $c = -\frac{cov(X,Y)}{var(Y)}$.

ונגדיר $\tilde{X} = \bar{X} + c(\bar{Y} - E[Y])$. נשים לב שאם $\bar{Y}$ יהיה גבוה מהתוחלת שלו, המקדם של c יהיה חיובי ולכן $\tilde{X} < \bar{X}$ כלומר, גם $\bar{X}$ יהיה יותר גבוה מהתוחלת שלו, באופן דומה ל־$\bar{Y}$ ולכן נוריד אותו מטה.

מאחר ולרוב גם $cov(X, Y)$ לא ידועה אנליטית, נוסיף שלב 0 לאלגוריתם:

0. נייצר M דגימות של X, Y עם משתנים מקריים משותפים $x_1, ..., x_m$ וגם $y_1, y_2, ..., y_m$.
בעזרת דגימות אלה נעריך את הפרמטר c על־ידי:

$cov(X, Y) \approx \frac{1}{m-2} \sum_{i=1}^{m} (x_i - \bar{x})(y_i - \bar{y})$.

את $var(Y)$ נחשב גם על סמך הדגימות הללו. בסוף שלב זה לא נשתמש יותר ב־M הדגימות אלא נשתמש בM דגימות חדשות.

לאחר מכן נעבור לשלב 1 באלגוריתם עם הערכה לפרמטר c.

כלל אצבע לרוב נקצה 10% מהמשאבים העומדים לרשותנו להערכת c.

הסבר כעת נרצה להבין מדוע שיטה זו עובדת ומדוע השונות של $\tilde{X}$ תהיה קטנה יותר מהשונות של $\bar{X}$.

טענה $E[\tilde{X}] = E[X]$

הוכחה $E[\tilde{X}] = E[\bar{X} + c(\bar{Y} - E[Y]) = E[\bar{X}] + c(E[\bar{Y}] - E[Y])$

המעבר האחרון הוא מלינאריות התוחלת. נשים לב שבשלב זה נוכל להתייחס ל־c כאל קבוע מאחר ו"זרקנו" את הדגימות של c.

לפי חוק המספרים הגדולים, לכל Z מתקיים $E[\bar{Z}] = E[Z]$ ולכן $E[\bar{Y}] = E[Y]$ ומאחר והביטוי האחרון מתאפס נקבל את השוויון $E[\tilde{X}] = E[\bar{X}] = E[X]$.

נחשב את השונות של X:

$var(\tilde{X}) = var(\bar{X}+c(\bar{Y}-E[Y])] = var(\bar{X})+2cov(\bar{X}, c(\bar{Y}-E[Y]))+var(c(\bar{Y}-E[Y])) =$

$= var(\bar{X})+2cov(\bar{X},\bar{Y})+c^2var(\bar{Y}) = var(\bar{X})+2cov(\frac{1}{m}\sum_{i=1}^m x_i, \frac{1}{m}\sum_{i=1}^m y_i)+$
$c^2var(\frac{1}{m}\sum_{i=1}^m y_i)$

$\bar{X}$ הוא ממוצע של M עותקים בלתי תלויים של X ו־ $\bar{Y}$ הוא ממוצע של M עותקים בלתי תלויים של Y.

ידוע ש־x_i, y_i מיוצרים עם אותם משתנים מקריים (כלומר, יש ביניהם קורלציה) ועבור $i \neq j$ x_i, x_j מיוצרים עם משתנים מקריים שונים ולכן הם בלתי תלויים.

$2cov(\frac{1}{m}\sum_{i=1}^m x_i, \frac{1}{m}\sum_{i=1}^m y_i) = \frac{1}{m^2}\sum_{i=1}^m\sum_{j=1}^m cov(x_i,y_j) = \frac{1}{m^2}m\cdot cov(X,Y) =$
$\frac{1}{m}cov(X,Y)$

נציב במשוואה ונקבל:

$$var(\frac{1}{m}\sum_{i=1}^m y_i) = \frac{1}{m^2}\sum_{i=1}^m var(y_i) = \frac{var(Y)}{m}.$$

שיעור 7

שיטת משתנה בקרה להורדת השונות במקום לחשב את $\bar{X}$ כקירוב ל$E[X]$ נחשב את $\tilde{X} = \bar{X} + c(\bar{Y} - E[Y])$ עבור $c = -\frac{cov(x,y)}{var(y)}$ ו־Y משתנה מקרי כך ש$E[Y]$ ידועה אנליטית.

באופן כללי, $var(\bar{X}) = \frac{1}{M}var(X)$ ולכן ההערכה לשגיאה הסטוכסטית היא $\frac{\sqrt{var(X)}}{\sqrt{M}}$.

בשיעור שעבר רצינו להראות ש־$v(\tilde{X}) < v(\bar{X})$.

$$v(\tilde{X}) = var(\bar{X}) + 2c\frac{cov(X,Y)}{M} + c^2\frac{1}{M}var(Y) =$$
$$var(\bar{X}) + 2\frac{-cov(X,Y)}{var(Y)}\frac{cov(X,Y)}{M} + \frac{(cov(X,Y))^2}{var(Y)^2}\frac{1}{M}var(Y) =$$
$$var(\bar{X}) - \frac{1}{M}(2\frac{cov^2(X,Y)}{var(Y)} - \frac{cov^2(X,Y)}{var(Y)}) = var(\bar{X}) - \frac{1}{M}\frac{cov^2(X,Y)}{var(Y)} < var(\bar{Y})$$

מאחר והביטוי $\frac{1}{M}\frac{cov^2(X,Y)}{var(Y)} > 0$, השונות אכן יורדת בחישוב $\tilde{X}$.

הערה נמשיך לחשב (נרצה לדעת בכמה השונות יורדת)

$$var(\tilde{X}) = var(\bar{X}) - \frac{1}{M}\frac{cov^2(X,Y)}{var(Y)} =$$
$$= \frac{1}{M}var(X) - \frac{1}{M}\frac{cov^2(X,Y)}{var(Y)} =$$
$$= \frac{1}{M}var(X)(1 - \frac{cov^2(X,Y)}{var(X)var(Y)}) =$$
$$= \frac{1}{M}var(X)(1 - \rho_{X,Y}^2)$$

כלומר, המקדם שבו השונות יורדת הוא $(1-\rho^2_{X,Y})$- ולכן, ככל ש- $\rho^2_{X,Y}$ יהיה גדול יותר, הורדת השונות תגדל.

יתרונות וחסרות של שיטת משתנה בקרה

יתרונות 1. השיטה קלה ליישום. היא דורשת משתנה מקרי Y המתואם במידה מסוימת עם X נתון. במציאות, לאופציות ונילה יש מתאם עם הרבה מאוד דברים.

2. השיטת תמיד עובדת. למעשה, ברגע שנוסיף משתנה בקרה לכל הפחות לא עשינו נזק (השונות תהיה קטנה/שווה לשונות).

3. השיטה לא דורשת הרבה משאבים נוספים מעבר לשיטה הרגילה (לתמחור פשוט). הוספנו את שלב 0, שדרש בסביבות ה- 10% משאבים נוספים וקיבלנו ירידה בשונות.

חסרונות 1. הירידה בשונות לא משמעותית.

$Importance\ sampling$

נרצה לתמחר את הפונקציה $h(X)$, למשל במקרה בו X הוא מחיר נכס הבסיס בפקיעה ו-h הוא תשלום האופציה בהינתן אותו מחיר נכס בסיס. ונניח ש-X הוא משתנה מקרי רציף.

$$E[h(X)] = \int_{-\infty}^{\infty} h(t) f_X(t)dt =$$

נכפיל ונחלק ב-$f_Y(t)$, עבור Y משתנה מקרי המתאר מחיר נכס בסיס אחר,

$$= \int_{-\infty}^{\infty} h(t)\frac{f_X(t)}{f_Y(t)} f_Y(t)dt = \int_{-\infty}^{\infty} h*(t) f_Y(t)dt = E[h*(Y)]$$

כרגע שכחנו מX ונתייחס למשתנה המקרי החדש Y.

במקום לחשב את $h(\bar{X})$ - ניקח סימולציה של $X_1, X_2, ..., X_n$, נציב ב-h ונקבל $h(X_1), h(X_2), ..., h(X_n)$ ונחשב ממוצע, $h(\bar{X}) = \frac{\sum_{i=1}^{n} h(X_i)}{M}$.

על מנת לחשב את $h*(Y)$, ניקח סימולציה של $Y_1, Y_2, ..., Y_M$, נציב ב-$h*$, נקבל $h*(Y_1), ..., h*(Y_M)$ ונחשב ממוצע, $h*\bar{(Y)} = \frac{\sum_{i=1}^{M} h(Y_i)}{M}$.

נרצה לעשות זאת כאשר $var(h*(Y)) < var(h(X))$ כך שנוכל לקבל שונות נמוכה יותר.

מתי אי השוויון הנ"ל יתקיים? כלומר, כיצד נבחר את Y כך ש- $var(h*(Y)) - var(h(X)) < 0$?

נפרק את הביטוי $var(h*(Y)) - var(h(X))$:

$$v(h*(Y)) - v(h(X)) = E[h*(X)^2] - E[h*(Y)]^2 - E[h(X)^2] + (E[h(X)])^2 =$$
$$= \int_{-\infty}^{\infty} h*(y)^2 f_Y(t)dt - \int_{-\infty}^{\infty} h(t)^2 f_X(t)dt =$$
$$= \int_{-\infty}^{\infty} h(t)^2 \frac{f_X(t)^2}{f_Y(t)^2} f_Y(t) = h(t)^2 f_X(t)dt$$
$$h*(s) = h(s)\frac{f_X(s)}{f_Y(s)} = \int_{-\infty}^{\infty} h(t)^2 f_X(t)\left(\frac{f_X(t)}{f_Y(t)} - 1\right)dt$$

נסתכל כעת על הביטוי האחרון. הביטוי $h(t)^2 f_X(t)$ הוא תמיד חיובי, הביטוי $\left(\frac{f_X(t)}{f_Y(t)} - 1\right)$ לפעמים שלילי ולפעמים חיובי. נָשים לב כי לא ייתכן שהביטוי שלילי בלבד, כי $\int f_X = \int f_Y = 1$.

הביטוי יהיה שלילי כאשר $\frac{f_X(t)}{f_Y(t)} - 1 < 0$, כלומר, כאשר $f_X(t) < f_Y(t)$ ויהיה חיובי במקרה ההפוך $f_X(t) > f_Y(t)$.

מכך נוכל להסיק שבכדי שהפרש השונויות יהיה שלילי, יש לבחור Y כך שהאינטגרל האחרון שלילי. אינטגרל זה יהיה שלילי אם נבחר את Y כך ש־

1. במקרה הראשון: $f_X(t) < f_Y(t)$ כאשר $h(t)^2 f_X(t)$ קרוב למקסימום שלו (גדול יחסית).
2. המקרה השני יהיה בדיוק ההפך־ $f_Y(t) < f_X(t)$ כאשר $h(t)^2 f_X(t)$ קטן יחסית.

דוגמא חשיבות הנקודה t (שווי נכס הבסיס בפקיעה) אשר בה האופציה מחוץ לכסף היא אפס. כלומר, אם אפשר להימנע ולייצר רק מסלולים שמכניסים אופציה של כסף, אני ארוויח בתמחור. כל מסלול בו הכסף הוא אפס לא רלוונטי עבורי.

לפיכך, אם נבחר נכס בסיס אחר Y הנמנע מנקודות בהן האופציה מחוץ לכסף, נוכל להוריד את שונות התמחור.

אלגוריתם כללי 1. נייצר מסלולים של Y במקום של X, $y_1, y_2, ..., y_M$.

2. לכל מסלול נחשב את היחס $LR_i = \frac{f_X(y_i)}{f_Y(y_I)}$ (יחס הנראות). המספר יהיה קטן עבור מסלולים נדירים.

3.נחשב את שווי תשלום האופציה בכל אחד מהמסלולים $h(y_i)$.

4. נכפול בLR ונקבל: $h*(y_i) = h(y_i) \cdot LR_i$. כאשר $h(y_i)$ הוא תשלום האופציה וLR_i הוא משקל המסלול ה־i.

5. נחשב ממוצע: $\bar{h*}(Y)=\frac{\sum h*(y_i)}{M}=\frac{1}{M}(\sum h(y_i)\cdot LR_i)$. הביטוי האחרון הוא ממוצע משוקלל והוא למעשה האומד ־$E[h(X)]$.

דוגמא נרצה לתמחר אופציה רחוק מחוץ לכסף, כלומר $s_0 << k$.

בשלב הראשון, אופציית $call$ רגילה. בשלב השני, אופציית $call$אסיאתית (כאשר נכס הבסיס־ממוצע המסלול עד הפקיעה).

1. רוב המסלולים לא יכניסו את האופציה לכסף כי דרושה עלייה גדולה במחיר נכס הבסיס. אם לא נשתמש בשיטה להורדת שונות, התמחור יהיה 0. תמחור בעזרת $sampling$ $importance$ ייתן תוצאה מדויקת יותר (קרוב ל0).

על מנת לדגום מסלולים המכניסים את האופציה לכסף, נרצה לדגום מסלולים עולים כך שבמקום $Z\sim N(0,1)$ נדגום $Z\sim N(\mu,1)$ כאשר μ הוא ערך חיובי. כתוצאה מכך, המסלולים שנקבל יהיו בעלי רעש חיובי מהרגיל ולכן יעלו.

ההחלפה תהיה של המשתנים $X\sim N(0,1),Y\sim N(\mu,1)$ במשתנים $f_X(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}},f_Y(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(t-\mu)^2}$.

ה־LR יהיה $e^{-t\mu+\frac{1}{2}\mu^2}=e^{-\frac{t^2}{2}+\frac{1}{2}t^2-\frac{1}{2}2t\mu+\frac{1}{2}\mu^2}=e^{-\frac{t^2}{2}+\frac{1}{2}(t-\mu)^2}=\frac{f_X(t)}{f_Y(y)}$.

בt, עבור כל מסלול, נציב את הערך שדגמנו:

דוגמות $y_i\sim N(\mu,1)$ במקום $x_i\sim N(0,1)$. הדגימה של המשתנה המקרי הנורמלי נכנסת לתוך ה־LR ולכן נחשב עבור המסלול ה־i את ה־LR_i, $LR_i=e^{-y_i\mu+\frac{1}{2}\mu^2}$.

תשלום $call$ אסייתית הוא $(\sum_{i=1}^N s_{ih}-k)_+$, כאשר $\sum_{i=1}^N s_{ih}$ הוא ממוצע על המסלול.

החלפנו במסלול ה־i, N משתנים מקריים $x_i\sim N(0,1)$ בN משתנים מקריים $y_i\sim N(\mu,1)$. כל החלפה כזו מכפילה את ה־LR ב־$e^{-y_i\mu+\frac{1}{2}\mu^2}$.

סה"כ ה־LR הוא $LR_i=e^{N\frac{1}{2}\mu^2-\mu\sum_{j=1}^N y_i}=e^{-y_1\mu+\frac{1}{2}\mu^2}e^{-y_1\mu+\frac{1}{2}\mu^2}...e^{-y_N\mu+\frac{1}{2}\mu^2}$.

באופן כללי, כאשר נחליף N משתנים מקריים x_i ב־y_i, הנוסחה ל־LR היא:
$\frac{f_{X_1}(y_1)}{f_{Y_1}(y_1)}\cdots\frac{f_{X_n}(y_n)}{f_{Y_n}(y_n)}$.

שיעור 8

Importance sampling

במקום X נדגום את $Y(\{y_i\})$ ונחשב את הממוצע של $g(y_i)\frac{f_X(y_i)}{f_Y(y_i)}$ כאומד לערך של $E[g(x)]$.

אם נחליף את $x^{(1)}$ ב-$y^{(1)}$, $x^{(2)}$ ב-$y^{(2)}$ וכן הלאה, נחשב ממוצע של $g(y_i^{(1)}, y_i^{(2)}, ..., y_i^{(n)})\cdot$ $\Pi_{k=1}^{n}\frac{f_{X^{(k)}}(y_i^{(k)})}{f_{Y^{(k)}}(y_i^{(k)})}$ כאומד לערך של $E[g(x^{(1)}, x^{(2)}, ...)]$.

דוגמא (*Importance sampling asian*) אנחנו רוצות לתמחר אופציה אסייתיות "עמוק מחוץ לכסף", ונרצה לדגום מסלולים המכניסים את האופציה לכסף, כלומר, מסלולים עולים, על מנת לייצר מסלול של GBM. נדגום הרבה משתנים מקריים נורמלים סטנדרטיים, $N(0,1)$, נחליף אותם ב-$N(\mu,1)$ כאשר $\mu > 0$ ונקבל מסלולים עולים. את שווי האופציה שנקבל עבור כל מסלול נכפול ב-LR - ההחלפות לפי החישוב הממוצע.

בחילוף אחד של $X \sim N(0,1)$ ב-$Y \sim N(\mu,1)$, נכפול ב- $LR= exp(-\mu y + \frac{\mu^2}{2})$.

כאשר נבצע הרבה חילופים, נכפול במכפלת ה-LR של כלל החילופים,

$$LR = exp(-\mu y_1 + \frac{\mu^2}{2})\cdot exp(-\mu y_2 + \frac{\mu^2}{2})\cdot ... \cdot exp(\mu y_n + \frac{\mu^2}{2}) =$$

והביטוי מקיים:

$$= exp(-\mu(y_1 + ... + y_n) + n\frac{\mu^2}{2})$$

שיטות להורדת שונות

תזכורת לפי חוק המגדל (או חוק התוחלת השלמה) מתקיים $E[E[X/Y]] = E[X]$.

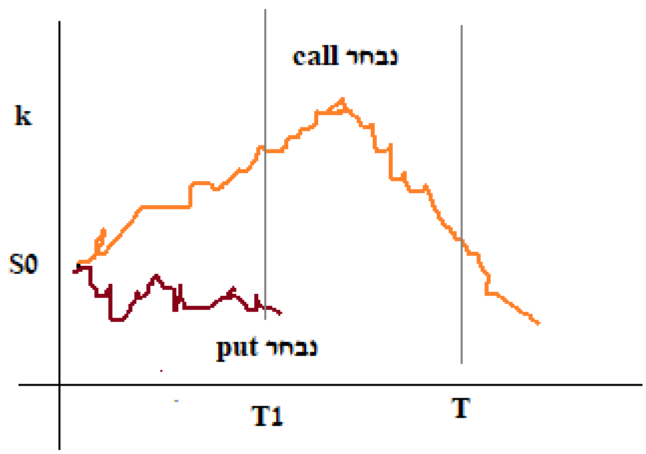
נשתמש בכלל זה על מנת להוריד שונות. ניזכר שנוסף לו קיים חוק השונות השלמה לפיו מתקיים: $var(X) = var(E[X/Y]) + E[var(X/Y)]$

בשתי השיטות הבאות שנראה נחשב בצורה אנליטית את אחת מהתוחלות בביטוי ואת השנייה נקרב בעזרת מונטה קרלו. בהתאם לכך, השונות של האומד שלנו תהיה אחת מהשונויות בפירוק, כאשר פירוק השונות הוא בהכרח פירוק של מספרים חיוביים ולכן כל אחת משונויות הפירוק תהיה קטנה יותר מ-$var(X)$. זה למעשה ייתן לנו את הירידה בשונות.

שיטה 3

שיטת התניה $Conditional\ monte\ carlo$ נניח כי לא ניתן לחשב אנליטית את $E[X]$ אבל כן ניתן לחשב אנליטית את $E[X/Y]$.

למשל אם נרצה לתמחר אופציית as you like it- אופציה אירופאית עם זמן פקיעה T (ומחיר התחלתי S_0, מחיר מימוש k) כך שצריך להחליט בזמן $T_1 < T$ האם היא אופציית $Call$ או אופציית Put (הפרמטרים זהים).



בדוגמא זו אנחנו רוצות לתמחר את האופציה.

$E[X]$הוא מחיר האופציה כאשר X הוא התשלום המהוון של האופציה.

Y הוא משתנה מקרי הקובע האם האופציה היא אופציית $Call$או אופציית Put . אם Y ידוע אזי, $E[X/Y = call] = Blsprice('C')$ או $E[X/Y = put] = Blsprice('P')$ונוכל לתמחר לפי בלק-שולס.

נבצע סימולציה ל-Y בלבד. נדגום את נכס הבסיס (GBM) בזמן T_1, S_{T_1}.

בנקודה זו נתמחר שתי אופציות, $Call$ וPut רגילות עבור פרק הזמן $T - T_1$ עם מחיר התחלתי S_{T_1} וסטרייק k. לפי נוסחת התמחור, נוכל לדעת בנקודה זו כמה כל אחת מבין האופצציות שווה (חישוב אנליטי). Y למעשה ייקבע לפי הערך הגבוה של האופציות כך שאם

שווי אופציית $Call$ גבוה יותר משווי Put, $Y = Call$ וההפך. נשים לב כי באותה נקודת זמן נוכל לקנות את האופציה הטובה ביותר ולכן נבחר את היקרה מהשניים. משלב זה, האופציה הופכת להיות אופציה רגילה, נהוון את הערך הגבוה מבין השניים לזמן $t = 0$ ונקבל דגימה של X/Y, שאותה נוכל לתמחר. לאחר ייצור של מספיק דגימות, נחשב ממוצע ונקבל אומד ל־$E[X]$.

דוגמאות 1. כל אופציה שבזמן $T_1 < T$ הופכת לאופציית ונילה תחת תנאי כלשהו.

2. אופציית מחסום, סימולציות עד פגיעה במחסום נתון ולאחר מכן תמחור בעזרת בלק־שולס.

יתרונות 1. לרוב השימוש בשיטה גורר ירידה בצריכת המשאבים.

2. תמיד יש ירידה בשונות.

חסרונות 1. החיסרון היחיד לשיטה הוא שהיא לא מאוד שימושית, כלומר, לא קיימות הרבה סיטואציות בהן ניתן להפעיל אותה.

הורדת השונות האומד ל$E[X]$ הוא $E[X/Y]$ (כאשר ניקח את ממוצע הדגימות של $E[X/Y]$). השונות של האומד היא $var(E[\bar{X/Y}]) = \frac{1}{n}var(E[X/Y]) \leq \frac{1}{n}var(X) = var(\bar{X})$, כאשר אי השוויון נכון לפי פירוק השונות ו־$var(\bar{X})$ היא השונות של האומד הסטנדרטי הממוצע. קיבלנו שהשונות תמיד יורדת.

שיטה 4 דגימה "מדורגת" Stratified sampling . באופן דומה, נסתכל על מגדל התוחלות $E_Y[E_X[X/Y]]$. אם Y הוא משתנה מקרי פשוט (בדיד) , אז מתקיים: $E_Y[E_X[X/Y]] = \sum_k E[X/Y=k]\cdot p_Y(k)$, כאשר $p_Y(k)$ ידועה אנליטית ו$E[X/Y=k]$ לא ידועה אנליטית. הסכום על הפרמטר k הוא על כל הערכים שY יכול לקבל. מאחר והגורם $E[X/Y=k]$ לא ידועה אנליטית, לכל ערך ש־Y יכול לקבל, נבצע סימולציה בנפרד. לאחר מכן נוכל לשלב את הסימולציות על־ידי $E_Y[E_X[X/Y]] = \sum_k E[X/Y=k]\cdot p_Y(k)$ יחד עם הידע האנליטי בנוגע ל־$p_Y(k)$.

ערכי Y מחלקים את הדגימה למספר שכבות, ומכאן שמה של השיטה. נרצה להסתכל על מקרה אחד מאוד חשוב של השיטה הזו.

מקרה פרטי חשוב נסמן $Y = I_A$ אינדיקטור של מאורע A.
$Y = \left\{ \begin{array}{ll} 1 & A, happens \\ 0 & Otherwise \end{array} \right\}$. יש שתי אפשרויות, הראשונה היא שהמאורע התרחש והשנייה היא המשלים, המאורע לא התרחש.

במקרה זה הנוסחה $E_Y[E_X[X/Y]] = \sum_k E[X/Y=k] \cdot p_Y(k)$ תהיה:
$E[X/Y=1]p_Y(1) + E[X/Y=0]p_Y(0) = E[X/A] \cdot p(A) + E[X/A^c](1-p(A)$.

מקרה פרטי עוד יותר A הוא מאורע כך שאם הוא לא מתרחש, אז $X = 0$. למשל, נוכל לחשוב על A כעל המאורע = "האופציה בכסף בזמן הפקיעה". במקרה זה, ללא תלות בתנאי האופציה, אם האופציה בכסף היא לא משלמת, וזה יישאר כך גם אם נהוון אחורה.

במקרה זה, הנוסחה $E_Y[E_X[X/Y]] = \sum_k E[X/Y=k] \cdot p_Y(k)$ שבעזרתה נתמחר, תהיה:

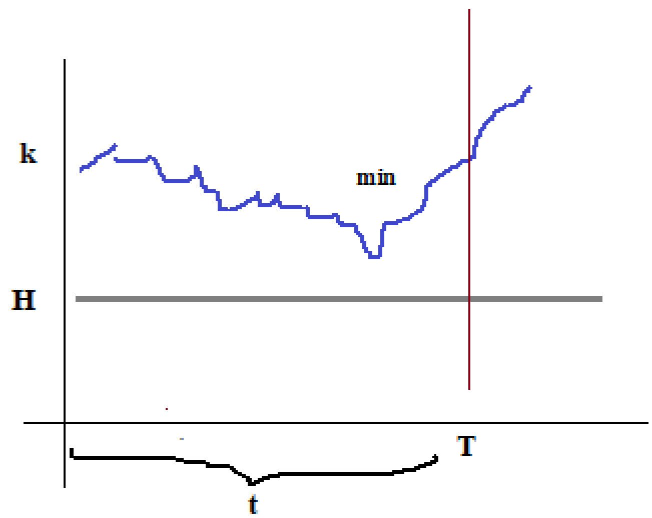
$E[X/A]p(A) + E[X/A^c]p(A^c) = E[X/A]p(A)$ כאשר $E[X/A^c] = 0$ (נובע מכך שבהינתן A^c, $X = 0$).

כלומר, יש לחשב $E[X] \approx E[X/A]p(A)$, כך ש־ $p(A)$ מחושב אנליטית , ו$E[X/A]$ מחושב על־ידי סימולציות.

דוגמא נרצה לתמחר אופציית מחסום על GBM, $down\ and\ out$. אם מחיר הנכס בין 0 ובין T יורד מתחת למחסום H אז האופציה בטלה.

הפרמטרים יהיו כרגיל־ S_0, k, T אליהם נוסיף פרמטר חדש H ־ המחסום.

האופציה משלמת בזמן T $(k - S_T) \cdot (k > S_T) \cdot (min_{0<t<T} S_T > H)$ כאשר שני הגורמים האחרונים במכפלה הם גורמים לוגיים (מקבלים 1 אם אי השוויון מתקיים ו־0 אחרת).



נשתמש ב־$Stratified\ sampling$ בהינתן המאורע $A=\{k>S_T\}$, כלומר, A הוא המאורע של "האופציה בכסף", והאופציה תהיה 0 אם A לא יתרחש.

חישוב $p(A)$:

$S_t \sim GBM(\mu,\sigma)$ ולכן $S_T \sim LN$ (לוג נורמלי). $S_T = exp(\mu - \frac{\sigma^2}{2}T + \sigma W_T)$ כאשר $W_T \sim N(0,T)$.

$P(A)$ היא למעשה $p(x>S_T)$, כך שנוכל להציג ולקבל:

$p(k>S_T) = p(k>exp((\mu-\frac{\sigma^2}{2})T+\sigma W_T)) = p(ln(k)>ln(S_0)+(\mu-\frac{\sigma^2}{2})T+\sigma W_T)) =$

נעביר הכל אגף חוץ מW_T ונקבל

$= p(ln(k)-ln(S_0)-(\mu-\frac{\sigma^2}{2})T>\sigma W_T)) =$

אנחנו יודעים ש־$\sigma W_T \sim N(0,\sigma^2 T)$ ולכן נתקנן,נחלק ב־$\sigma\sqrt{T}$ ונקבל $N(0,1)$,

$= p((ln(k)-ln(S_0)-(\mu-\frac{\sigma^2}{2}))/\sigma\sqrt{T}>Z) = \Phi((ln(k)-ln(S_0)-(\mu-\frac{\sigma^2}{2})/\sigma\sqrt{T})$

לאחר חישוב $p(A)$, נרצה לדגום את X/A וזאת נעשה בעזרת הגשר הבראוני.

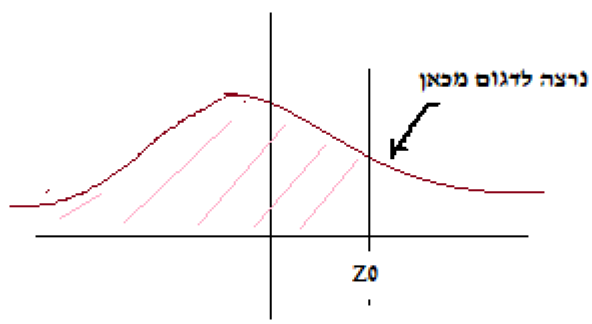
שיעור 9

Stratified sampling

בהינתן מאורע A כך ש־$X=0$ כאשר A לא מתרחש. אזי, $E[X]=E[X/A]p(A)$.
כפי שראינו בשיעור הקודם, את $p(A)$ נחשב ואת $E[X/A]$ נסמלץ.

דוגמא בדוגמה בשיעור שעבר ראינו X ־ התשלום של אופציית מחסום כאשר A הוא האופציה בכסף במועד הפקיעה, $A=\{S_T<k\}$.
בשיעור שעבר חישבנו את $p(A)$ וכעת נייצר דגימות של X/A (כלומר, נדגום את תשלום האופציה בהינתן שהיא בכסף).

האלגוריתם 1. נדגום את נכס הבסיס בזמן T בהינתן $S_T<k$. שאלה זו שקולה לדגימת $Z/Z<Z_0$ כאשר $Z\sim N(0,1)$.
בשבוע שעבר ראינו שהמאורע A, $\{S_T<k\}$. שקול למאורע $\{Z<\frac{ln(\frac{k}{S_0})-(r-\frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\}$
נסמן $Z_0\frac{ln(\frac{k}{S_0})-(r-\frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$ ואכן $S_T<k \iff Z<Z_0$.
באופן כללי, כשרוצות לדגום $Z/Z<Z_0$,



נשתמש בשיטת ההיפוך: ההצטברות של Z, $\Phi(t)$ ידועה ובפייתון ניתן לחשב את $\Phi^{-1}(y)$.
ההצטברות של המשתנה המקרי $W=Z/Z<Z_0$:

$$F_W(t)=\left\{\begin{array}{ll}\frac{\Phi(t)}{\Phi(Z_0)} & t<Z_0\\ & t\geq Z_{01}\end{array}\right\}$$

$$y=\frac{\Phi(t)}{\Phi(Z_0)}$$

$t = \Phi^{-1}(\Phi(Z_0)y)$

ולכן, $F_W^{-1}(y) = \Phi^{-1}(y\Phi(Z_0))$ והחשיבות הגדולה היא שאת $F_X^{-1}(y)$ ניתן לחשב.

סיכום (שלב מספר 1) 1. דוגמות $u \sim U[0,1]$.

2. מציבות ב־F_W^{-1} ומקבלות $w = \Phi^{-1}(u \cdot \Phi(Z_0))$ כאשר u היא דגימה של $Z/Z < Z_0$.

3. מציבות את w בנוסחה של GBM, $S_T = S_0 exp((r - \frac{\sigma^2}{2})T + \sigma\sqrt{T}w)$.

החלק השני $\sigma\sqrt{T}w$ הוא התנועה הבראונית W_T, כך שנקבל דגימה אחת של $S_T/S_T <$ k.

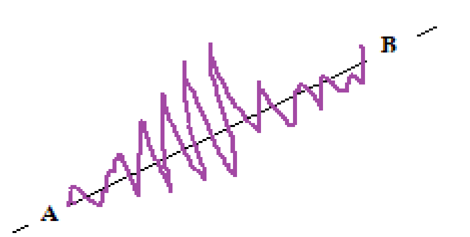
שלב 2 נרצה לדגום את שאר המסלול S_t בהינתן S_0 ו־S_T. כלומר, כעת אנחנו יודעות את תחילת המסלול S_0 ואת סופו S_T ונרצה להשלים את המסלול. לשם כך נעזר במשפט הגשר הבראוני. המשפט מאפשר לנו, בהינתן קטע $[a,b]$ ו־W_a, W_b, לדעת את ההתפלגות של W_t לכל $t \in [a,b]$.

משפט (הגשר הבראוני) בהינתן תנועה בראונית W_t ו־$a < b$ זמנים כלשהו. אזי, לכל $a < t < b$,
$$W_t/W_a = A, W_b = B \sim N\left(\frac{A(b-t)+B(t-a)}{b-a}, \frac{(t-a)(b-t)}{b-a}\right)$$

הפרמטר הראשון בהפלגות הנורמלית הוא μ התוחלת, והשני הוא השונות, σ^2. A, B הוא הלוגריתם של המחיר, הרעש. נשים לב כי ליחידות של W_t אין משמעות מבחינה זו.

הנוסחה של $W_t/W_a = A, W_b = B$ היא אינטרפולציה לינארית. אנחנו יודעות כי התנועה הבראונית עולה מ־A ל־B, ומאחר ואין העדפה לזמנים, נוכל להניח כי היא עולה בקצב קבוע.

האינטרפולציה לינארית של $(a, A), (b, B)$ עובדת באופן הבא־ אם נסתכל על האמצע, "עברנו" בממוצע חצי מהדרך. בנוגע לשונות, נוכל להסתכל עליה כעל מעין פרבולה בוכה עם נקודות קצה ב־0 כל שהיא מקבלת את המקסימום שלה באמצע הדרך. מסלול טיפוסי של גשר בראוני ניתן לתיאור באופן הבא,



כאשר נתחיל ב־A, נעלה ל־B, והרעש יגדל ככל שמתקרבים לאמצע ויקטן בצדדים.

הערה הנוסחה נכונה כל עוד לא ידוע דבר על אמצע W_t בקטע $[a,b]$ מלבד הנתונים על הקצוות $W_a = A, W_b = B$.

נרצה לייצר מסלול של W_t בהינתן W_0, W_T על־ידי שימוש במשפט. בשלב הראשון נחלק את הקטע $[0,T]$ ל־n חלקים שווים־ $0, h, 2h, ..., nh = T$ כאשר אנחנו יודעים את הערכים W_0 ו־$W_{nh} = W_T$ ונרצה לחשב את W_i לכל $i = kh$. כעת נרצה לנוע משמאל לימין על־גבי השלד, בדומה לשיטת אוילר. בהינתן W_{ih}, נחשב את $W_{(i+1)h}$ בעזרת הגשר הבראוני בין $(0, W_0), (T, W_T)$ כאשר $(ih, W_{ih}) = (a, W_a = A), (T, W_T) = (b, W_b = B)$ כך שבסימוני המשפט, $t = (i+1)h$.

נשים לב שתמיד נרצה להיות במצב בו לא נדע דבר על הקטע כך שכל מידי שיש בידינו נשאיר מאחור. W_T יישאר קבוע בכל אחד מהמקרים, אך נקודת ההתחלה תשתנה בכל זמן. נפעל באופן דומה עד שנגיע לנקודה האחרונה בה נחשב גשר בראוני עם $((n-2)h, W_{(n-2)h})$ ועם (T, W_T).

הערות

1. במעבר מהאיטרציה ה־i לאיטרציה ה־$i+1$ נחליף את הנקודה (a, A) ואת הזמן t בגשר הבראוני, $(ih, W_{ih}) \rightarrow ((i+1)h, W_{(i+1)h}) : (a, A)$ צד שמאל של הגשר.
2. (b, B) תמיד קבועה על (T, W_T)
3. על מנת לדגום $W_t \sim N(\mu_t, \sigma_t^2)$,

א. דוגמות $Z \sim N(0,1)$.

ב. מחשבות $\mu_t + \sigma_t Z$ כדגימה של W_t.

סיום שלב 2 בסוף הלולאה נקבל מסלול של W_t בהינתן מאורע A.

שלב 3 נציב את המסלול של W_t בנוסחה של GBM ונקבל מסלול של S_t:

$S_t = S_0 exp((r - \frac{\sigma^2}{2})t + W_t)$.

שלב 4 מתמחרות את האופציה כרגיל ומקבלות אומד ל$E[X/A]$.

שלב 5 כופלות בחישוב של $p(A)$, כפי שראינו קודם לכן, ומקבלות אומד ל$E[X]$ (מחיר האופציה) כאשר $E[X] \approx E[X/A]p(A)$.

הערה יישום האלגוריתם נמצא בקובץ $brownian\ bridge$.

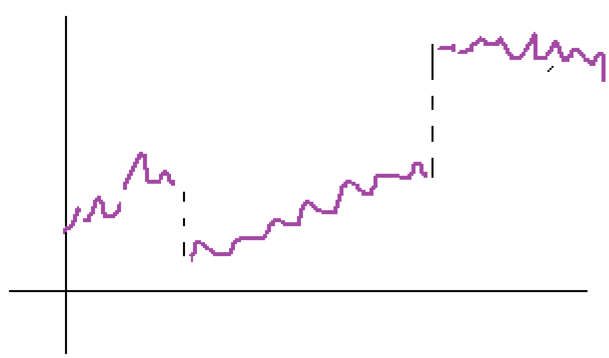
תהליכי קפיצה

בנושא זה נפתור ונייצר מסלולים של משוואה דיפרנציאלית סטוכסטית מהצורה הבאה:

$dS_t = S_t \cdot (\mu dt + \sigma dW_t + dJ_t)$.

נשים לב כי החלק $S_t(\mu dt + \sigma dW_t)$ מוכר לנו־ משוואות GBM. החלק $S_t dJ_t$ מהווה חידוש כאן והוא למעשה מייצג את הקפיצות בS.

נוכל לחשוב על מסלול של S_t באופן הבא:



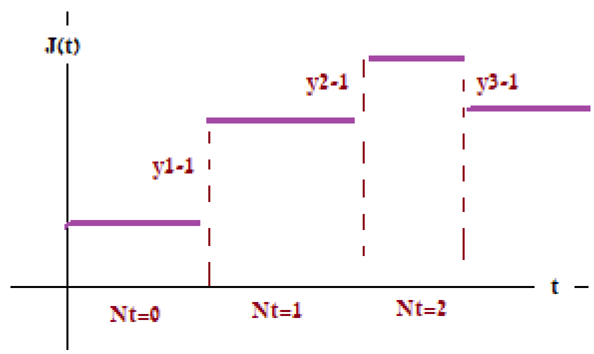
בין קפיצה לקפיצה הרכיב dJ_t של התהליך הוא 0, כלומר, נקבל GBM סטנדרטי.

התהליך J_t קובע את גודל הקפיצות ואת זמנן, ונחלק אותו לשני חלקים:

לרוב , $J_t = \sum_{i=1}^{N_t}(y_i - 1)$, כאשר N_t הוא תהליך פואסון- קובע את מועד הקפיצות ו-y_i הוא משתנה מקרי בעל התפלגות כלשהי- כאשר y_i, y_j לכל $i \neq j$ בלתי תלויים ושווי התפלגות.

בשלב הראשון נניח כי $y_i \sim logN(a, b^2)$, עבור a, b פרמטרים כלשהם.

גודל הקפיצה הi ב-J_t הוא $(y_i - 1)$.



במעבר בין $N = 0$ ובין $N = 1$ מתרחשת הקפיצה הראשונה, ובאופן דומה עבור הקפיצה השנייה וכן הלאה. גובה הקפיצה ה-i יהיה $y_i - 1$ כאשר נשים לב שקפיצה יכולה להיות שלילית, והיא למעשה תמיד ביחס ל-J הקודם.

בכל נקודה בין קפיצות ב-J, $dJ = 0$, ובנקודת קפיצה של J_t, יש קפיצה בS_t ואז $dS_t = S_t - S_{t^-}$ כאשר S_{t^-} הוא הגבול משמאל של S בנקודה t. באופן דומה, $dJ_t = J_t - J_{t^-}$, ונתייחס ל-dt, dW_t כאל אפס (לא קופצים).

כלומר גודל הקפיצה ב-S ייתן לנו את המשוואות:

$$S_t - S_{t^-} = S_{t^-}(J_t - J_{t^-}) = S_{t^-}(y_k - 1)$$

$J_t - J_{t^-}$ היא הקפיצה בJ בזמן t.

נעביר אגף, נפתח סוגריים ונקבל:

$$S_t = S_{t^-} + S_{t^-}y_k - S_{t^-}$$

מכאן כי $S_t = S_{t^-}y_k$. כלומר, הקפיצה בS_t היא כפלית. בנקודת הקפיצה S_t משתנה

פי y_k. נשים לי כי הכפליות מתאימה למודל השוק.

הפיתרון למד"ס הוא $S_t = S_0exp((\mu-\frac{\sigma^2}{2})t+\sigma W_t))\cdot\Pi_{k=1}^{N_t}y_k$, כאשר הרכיב הראשון הוא התנועה הבראונית הגאומטרית והרכיב השני הוא כלל הקפיצות הכפליות.

מכיוון שבחרנו y_k לוג נורמליים וגם GBM הוא לוג נורמלי - מכפלה של לוג נורמליים נשארת לוג נורמלית. כלומר- S_t עם הקפיצות הוא לוג נורמלי.

שיעור 10 (שישי)

תהליכי קפיצה תהליך קפיצה כללי הוא תהליך מהצורה $dS_t = a(S_t)dt + b(S_t)dW_t + c(S_t)dJ_t$ כאשר a, b, b הן פונקציות של S_t (לפעמים גם של t).

$c(S_t)dJ_t$ הוא חלק הקפיצה הגורם לתהליך להיות לא רציף.

ראינו ש־$J_t = \sum_{i=1}^{N(t)}(y_i-1)$, J_t הוא סכום כל הקפיצות עד זמן t (גודל הקפיצות). $N(t)$ הוא מספר הקפיצות עד זמן t- לרוב $N(t)$ יהיה תהליך פואסון, $N(t)\sim poiss(\lambda t)$ אם כי ישנם מקרים בהם $N(t)$ יהיה תהליך נקודתי אחר (לא בקורס). y_i-1 הם משתנים מקרים כלשהם, בלתי תלויים ושווי התפלגות. אם λ תלוי בתנודתיות, יהיו יותר קפיצות.

בין נקודת קפיצה אחת לשנייה התהליך קפיצה מתאפס כמו תהליך איטו רגיל ומקיים את המשוואה $dS_t = a(S_t)dt + b(S_t)dW_t$ (כלומר, $dJ = 0$). בנקודת הקפיצה ה־i, התהליך קופץ לפי המשוואה $dS_t = c(S_t)dJ_t$ (כלומר $dt = 0, dW_t = 0$) ומתקיים: $S_t - S_{t^-} = c(S_{t^-})(y_i-1)$, כאשר S_{t^-} הערך רגע לפני הקפיצה, וS_t הוא הערך אחרי הקפיצה. S_t מקיים: $S_t = S_{t^-} + c(S_{t^-})(y_i-1)$.

סימולציה של תהליך קפיצה נסמלץ תהליך קפיצה בין 0 ובין T. המסגרת היא תמחור אופציה על נכס בסיס S_t המהווה תהליך קפיצה. לשם כך יש לדגום מסלול של S_t.

בעזרת הרחבה של שיטת אוילר מריומה, נפעל לפי השלבים הבאים:

1. נחלק את הקטע $[0,T]$ לN קטעים קטנים באורך $h = \frac{T}{N}$.
2. נדגום $N(T)\sim poiss(\lambda T)$, N יהיה מספר הקפיצות בין 0 ובין T.
3. "נפזר" $N(T)$ קפיצות בין 0 ובין T. נעשה זאת באופן הבא- נדגום $N(T)$ ערכים אחידים בין 0 ובין T', $U[0,T]$ בלתי תלויים, $\tau_i\sim U[0,T]$ עבור $i = 1,2,...,N(T)$.
4. נדגום $N(T)$ ערכים מהתפלגות של y_i עבור $y_1, y_2, ..., y_{N(y)}$.

בשלבים 2,3,4 קיבלנו את כל המידע עבור מסלול של J_t. שלב 2 מתאר את מספר הקפיצות, שלב 3 מתאר את מיקום הקפיצות ושלב 4 מתאר את גודל הקפיצות.

5. נשבץ את הזמנים τ_i ברשת וניצור רשת חדשה, $0 = t_0 < h < \tau_1 < ... < t_N = T$.
הרשת החדשה מורכבת גם מהנקודות $0, h, 2h, 3h, ...Nh$ וגם מהנקודות $\tau_1, \tau_2, ...\tau_{N(T)}$ לפי סדר הופעתן בקטע.

נשים לב כי הרשת החדשה לא אחידה במרחקים בין 2 נקודות (נצטרך לקחת זאת בחשבון בהמשך).

6. בין נקודה לנקודה על הרשת (השלד) החדשה, נשתמש באוילר־מריומה,

$dS_t = a(S_t)dt + b(S_t)dW_t$.

$S_{t_i} = S_{t_{i-1}} + a(S_{t_{i-1}})(t_i - t_{i-1}) + b(S_{t_{i-1}})(W_{t_i} - W_{t_{i-1}})$.

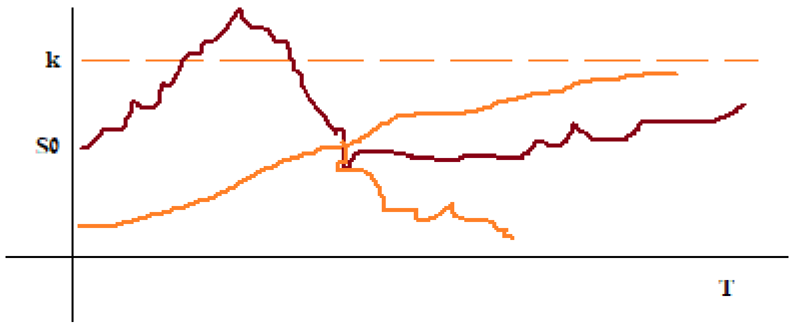
7. אם t_i היא לא נקודת קפיצה, נציב $S_{t_i^-} = S_{t_i}$ ונמשיך לt_{i+1}. אם t_i היא נקודת קפיצה, נניח נקודת הקפיצה ה־k, כלומר $t_i = \tau_k$, נציב: $S_{t_i} = S_{t_i^-} + c(S_{t_i^-})(y_k - 1)$ ונמשיך ל־t_{i+1}.

שיעור 11

תמחור אופציות אמריקאיות

אופציה אמריקאית ניתנת למימוש מיידי בכל זמן $0 < t < T$, כל שבכל t, נוכל לקבל את תשלום האופציה עבור $Vanilla\ put$, $(k - S_t)_+$ מייד.

הערה אופציה אמריקאית תמיד שווה יותר מהאופציה האירופאית המקבילה לה (ראינו ב"תמחור אופציות").



המוקדם Early Exercise boundry

קו המימוש המוקדם הוא קו כאשר ברגע שהמניה פוגעת בו, כדאי לממש את האופציה מוקדם (נניח שבמקרה זה המניה יורדת נמוך מספיק כך שמימוש מוקדם יהיה שווה יותר מהמשך ההחזקה של האופציה). הקו תלוי בהיסטוריה של S_t בקטע $[0,t]$, כלומר, עד נקודת הזמן הנוכחית.

אלגוריתמים לתמחור אופציות אמריקאיות

עץ בינומי נרצה לתמחר $put\ vanila$ על GBM, $dS_t = S_t(rdt + \sigma dW_t)$.

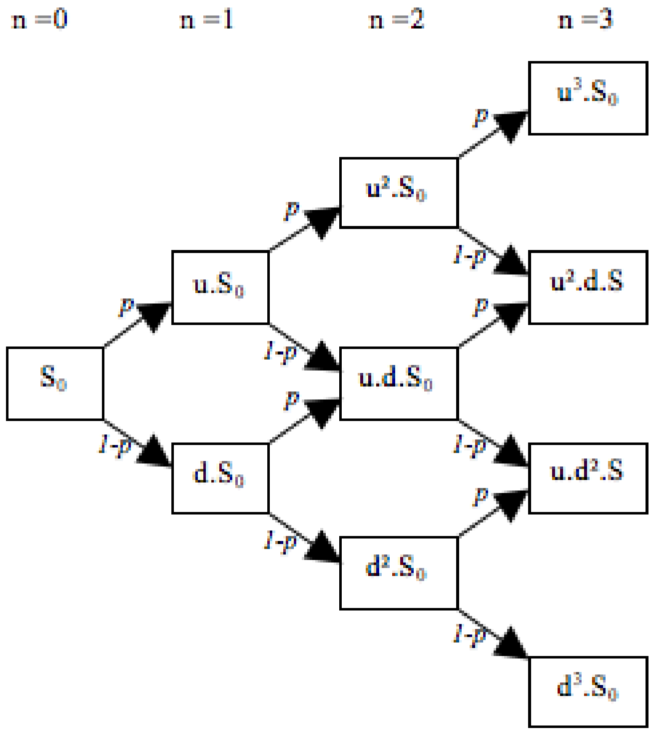
נחלק את הקטע $[0,T]$ לn קטעים באורך $h = \frac{T}{n}$.

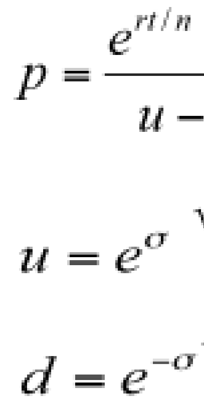
נניח כי בין כל שתי נקודות זמן עוקבות נכס הבסיס יכול לעלות פי u או לרדת פי d ונבחר $u = e^{\sigma\sqrt{h}}$, $d = e^{-\sigma\sqrt{h}}$.

כלומר, לS_{ih} יש 2 אפשרויות עבור $S_{(i+1)h}$. $S_{(i+1)h}$ תהיה uS_{ih} או תהיה dS_{ih}.

על מנת שקפיצה למעלה תבטל קפיצה קודמת, צריך להתקיים $u \cdot d = 1$ ־ במקרה זה הענפים של העץ מתחברים זה לזה (הנחה זו הופכת את העץ פשוט לעבודה).

על מנת שהמסלולים על העץ יהיו דומים למסלולים של GBM, נגדיר את ההסתברות לעלייה $p = \frac{e^{rd} - d}{u-d}$ ונבנה את העץ.





השווי יהיה $[k - []]$ כאשר [] מתאר את הביטוי הריבוע (שווי האופציה).

בהינתן העץ, נוכל לתמחר כל אופציה נתונה. ונרצה לתמחר אופציה אמריקאית על העץ מהסוף להתחלה:

1. נתחיל בזמן הפקיעה T. שווי האופציה בזמן זה היא התשלום של $(k - S_T)_+$.
2. עוברותרמה אחת אחורה ומחשבות שני ערכים:

א. שווי המימוש המיידי (כלומר, מה היה שווי האופציה לו הייתי מממשת אותה באותו רגע?). השווי הוא $(k - S_t)_+$ כאשר S_t הוא הערך של S בקודקוד הנוכחי. לכל קודקוד ברמה $(n-1)h$.

ב. שווי החזקת האופציה: התוחלת המהוונת של שווי האופציה בשני הקודקודים הבנים של הקודקוד הנוכחי (מהוון): אם נסמן את שווי האופציה במקרה של עלייה בV_{up} ואת שווי האופציה במקרה של ירידה V_{down} נקבל $e^{-rh}(pV_{up} + (1-p)V_{down})$.

3. אם שווי המימוש המיידי עולה על שווי ההחזקה־ כדאי לממש ולכן שווי האופציה במקרה זה הוא $(k-S_t)_+$. אחרת, אם שווי המימוש המיידי קטן משווי ההחזקה, לא נממש ולכן שווי האופציה יהיה שווי ההחזקה. $V_{back}=max((k-S_t)_+,e^{-rh}(pV_{up}+(1-p)V_{down}))$.

השווי בקודקוד הנוכחי יהיה המקסימום של שווי מימוש מיידי ושווי ההחזקה.

4. חוזרות באופן איטרטיבי על החישוב של 2 ו־3 מחשבות את שווי האופציה ברמה i בעזרת הרמה ה־$i+1$.

5. ברמה ה־0 החישוב יניב את שווי האופציה בהווה.

יתרון האלגוריתם מהיר הרבה יותר משאר האלגוריתמים שראינו.

חסרון האלגוריתם לא עובד כאשר נכס הבסיס הוא צירוף של מספר מניות (כל נכס שלא מתנהג כ־GBM), למשל, $S_t=max(S_t^{(1)},S_t^{(2)})$.

כל שאר האלגוריתמים לתמחור אופציות אמריקאיות משתמשים בעיקרון המופיע בשיטת העץ הבינומי: תכנון דינמי כך שעל מנת לתמחר את האופציה בזמן $t=0$, יש לפתור בעיות קטנות יותר: תמחור בזמן ih בהינתן S_{ih}. שני השלבים של התכנון הדינמי הם:

1. פירוק הבעיה לבעיות קטנות התלויות זו בזו.
2. פיתרון הבעיות מהסוף להתחלה.

מעתה והלאה נניח הנחה מקלה: את האופציה אפשר לממש רק במספר סופי של זמנים ידועים מראש (למשל סוף כל יום מסחר). לפיכך, כאשר ניקח מספר זמני מימוש גדול,נקבל קירוב לאופציה האמריקאית.

אופציה מסוג זה נקראת אופציה ברמודיות (בין אירופה לאמריקה).

תמחור אופצית $put\ vanila$ **ברמודית במסגרת תכנון דינמי** נניח שזמני המימוש הם $0=t_0<t_1<...<t_n=T$. נרצה לחשב את V_0.

לכל t_i, נסמן את שווי האופציה בזמן t_i ב־V_{t_i}. מתחילות מהסוף ומחשבות את V_{t_i} בעזרת $V_{t_{i+1}}$.

1. בזמן הפקיעה $V_T=(k-S_T)_+$.
2. במועד המימוש הראשון לפני הפקיעה מחשבות שני ערכים:

א. שווי המימוש המיידי $(k-S_{t_{n-1}})_+$.

ב. שווי החזקה: $H_{t_{n-1}}=e^{-r(T-t_{n-1})}\cdot E[V_T/S_{t_{n-1}}]$. היוון תוחלת שווי האופציה במועד המימוש הבא בהינתן הערך הנוכחי של נכס הבסיס.

3. שווי האופציה בזמן t_{n-1} הוא הגדול מבין השניים, $V_{t_{n-1}}=max((k-S_{t_{n-1}})_+,H_{t_{n-1}})$.

4. חוזרות על 2־3 באופן איטרטיבי מהסוף להתחלה. נוסחת האיטרציה היא:

$V_{t_{i-1}}=max\left((k-S_{t_{i-1}})_+,H_{t_{i-1}}\right)$.

$H_{t_{i-1}}=e^{-r(t_i-t_{i-1})}E[v_i/S_{t_{i-1}}]$.

5. לבסוף נקבל את V_0־ מחיר האופציה בהווה.

הערה ההבדל בין האלגוריתמים השונים הוא גישות שונות לקירוב הערך $E[v_i/S_{t_{i-1}}]$. דמיון נוסף בין האלגוריתמים הבאים לתמחור אופציות אמריקאיות הוא שלא ניתן למצוא אומד בלתי מוטה ל־V_0: מוצאות שני אומדים־ אחד מוטה למעלה ואחד מוטה מטה כך ש: $E[V_L]<V_0<E[V^H]$. V_L האומד התחתון, V^H האומד העליון, V_0 הוא הערך שרוצים לחשב.

שיטת העץ המקרי $Random\ Tree$ הרעיון הוא ייור עץ מסלולים של נכס הבסיס כאשר S_0 יכול להתקדם למספר כיוונים, $S_{t_1}^1,S_{t_{12}}^2,...,S_{t_1}^{n_1}$.

בהמשך $S_{t_1}^1$ מתפצל ל־$S_{t_2}^{11},...,S_{t_2}^{1n_2}$ וכן הלאה. דוגמות מסלול מ $S_{t_1}^1$ ל־ S_{t_2}.

עבור כל קודקוד נייצא מסלולים עד לt_2, כאשר המסלולים לא מתאחדים. במקרה בו יש מספר גבוה של זמני מימוש, האלגוריתם יקר מידי (האלגוריתם אקספוננציאלי בזמן הריצה ביחס למספר זמני המימוש).

שיעור 12

אלגוריתם העץ המקרי $Random\ Tree$ **לתמחור אופציות ברמודיות** נרצה לתמחר אופציית $put\ vanilla$ עם זמני מימוש אפשריים, $0=t_0<t_1<t_2<...<t_n=T$ על נכס בסיס S_t.

תנאי בסיס ניתן לדגום מסלולים של S_t בשיטה כלשהי.

האלגוריתם שלב 0: בוחרות "קבוע פיצול" b , כך שמכל ענף בעץ נייצר עוד b ענפים חדשים ברמה הבאה. אותו b יישאר קבוע על פני כלל הענפים בעץ.

שלב 1: בניית העץ. נתחיל מ־S_0 בזמן t_0, ממנו נוציא b ענפים בזמן t_1, ונסמנן ב־ $S_{t_1}^1, S_{t_1}^2, ..., S_{t_1}^b$.

נדגום את S_t בזמן t_1 בהינתן הערך של S_0 . ללא תלות בשיטה, נרצה לקבל b דגימות שיתחילו בזמן t_0. נדגום באותה שיטה את אותן דגימות ונסמן אותן ב־ $S_{t_1}^1, ..., S_{t_1}^b$.

שלב 2: עבור כל דגימה, $S_{t_1}^i$ מהשלב הקודם, נדגום את S_{t_2}בהינתן הערך בזמן t_1, $S_{t_1}^i$ (הענפים הללו הם המשך המסלול) ונסמן את הדגימות החדשות ב־ $S_{t_2}^{i_1}, ...S_{t_2}^{ib}$.

נשים לב שכל הדגימות החדשות הן בלתי תלויות.

אם למשל היינו צריכות לדגום את $S_{t_2}^{ij}$ בעזרת הנוסחה האנליטית של GBM, היינו משתמשות בנוסחה $S_{t_2} \sim S_{t_1}^i \cdot exp((r - \frac{\sigma^2}{2})(t_2 - t_1) + \sigma(W_{t_2} - W_{t_1}))$.

$S_{t_1}^i$ היא נקודת ההתחלה ומתקיים $W_{t_2} - W_{t_1} \sim N(0, t_2 - t_1)$.

שלב 3: חוזרות באופן איטרטיבי על שלב 2 עד הזמן האחרון $t_n = T$ כך שבסופו נקבל עץ ערכים מהצורה $S_{t_k}^{i_1 i_2 i_3, ..., i_k}$. האינדקסים למעלה מתארים את ערכי הפניות והאינדקס למטה פירושו הרמה בה הערך נמצא.

$S_{t_k}^{i_1 i_2 i_3, ..., i_k}$ מתאר את הערך בזמן t_k המתקבל בעץ בענף ה־i_{k-1} על הענף ה־i_{k-2} וכן הלאה , של הענף ה־i_1. כלומר, הערך העליון מראה לנו אילו פניות נלקחו. למשל, $S_{t_4}^{1451}$ נמצא ברמה הרביעית ומגיעות אליו לאחר שבזמן 0 לקחנו את הענף הראשון, בזמן 1 לקחנו את הענף הרביעי, בזמן 2 לקחנו את הענף החמישי ובזמן 3 לקחנו את הענף הראשון.

בצורה יותר מדויקת, נוכל להסתכל על השלבים הבאים עד קבלת הערך:

1. $S_{t_1}^1$,
2. $S_{t_2}^{14}$,
3. $S_{t_3}^{145}$,
4. $S_{t_4}^{1451}$.

מציאת אומד עליון למחיר האופציה נמצא $\hat{V}$ כך ש־$V < E[\hat{V}]$ כאשר V הוא מחיר האופציה. הערכת $\hat{V}$ מתבצעת בדומה לתמחור האופציה על העץ הבינומי.

1. עבור על אחד מהקודקודים ברמה t_n (הרמה האחרונה) של העץ מחשבות את שווי האופציה בפקיעה: $V^{i_1 i_2 \dots i_n} = (k - S_{t_n}^{i_1 i_2 \dots i_n})_+$.

השווי של כל קודקוד הוא $(k-$ מחיר נכס הבסיס בקודקוד$)_+$. לאחר שיש לנו את השווי ברמה האחרונה, נלך רמה אחת אחורה.

2. בהינתן כל ערכי V_{t_n}, נעריך את ה־$Holding\ value$ (שווי ההחזקה) בקודקוד $i_1 i_2 \dots i_{n-1}$, ברמה t_{n-1} (יש $n-1$ אינדקסים) בתור הממוצע המהוון של שווי האופציה בכל הקודקודים הבנים שלו־ $H_{t_{n-1}}^{i_1 i_2 \dots i_{n-1}} = e^{-r(t_n - t_{n-1})} \cdot \frac{1}{b} \sum_{j=1}^{b} V_{t_n}^{i_1 i_2 \dots i_{n-1} j}$.

הסכימה היא על אינדקס j־ התנועה על העץ זהה להגעה לקודקוד שנרצה לחשב דרכו ואז עוברות על האפשרויות של הקודקודים הבנים שלו (האפשרויות של j).

3. מחשבות את השווי בקודקוד $i_1 i_2 \dots i_{n-1}$ במחיר המקסימום של ה־$Holding\ value$ וערך המימוש המיידי:

$V_{t_{n-1}}^{i_1 \dots i_{n-1}} = max\left(H_{t_{n-1}}^{i_1 i_2 \dots i_{n-1}}, (k - S_{t_{n-1}}^{i_1 \dots i_{n-1}})_+\right)$.

בבינומי יודעות בדיוק מה העץ, כלומר, יודעות שאין עוד אופציות במרחב המדגם. במקרה המתואר לעיל, לא ניתן לדעת את קודקודי העץ עד בניינו.

4. חוזרות באופן איטרטיבי על 2 ועל 3 ובזמן 0 מקבלות את האומד $\hat{V} = V_{t_0}$.

כאן $V < E[\hat{V}]$, בניגוד לעץ הבינומי.

כאן למעשה צמצמנו את נקודת המבט על העולם למספר קטן של מסלולים. העתיד יהיה בהכרח אחד מהמסלולים הללו ועל סמך זאת נוכל להחליט האם להמשיך להחזיק את האופציה. העובדה ש"הצצנו" לעתיד נתנה לנו יתרון לא הוגן ולכן המחיר האמיתי יהיה גבוה יותר מהאומד. בכל זמן מימוש אפשרי, ההחלטה האם לממש או לא מתקבלת על סמך שווי האופציה בהמשך העץ (S_t יכול לנוע בהרבה מסלולים החל מהזמן הנוכחי אבל אנחנו יודעות שהתנועה תהיה באחד מהמסלולים שדגמנו בעץ). השימוש בידע מהעתיד מוביל להחלטות אופטימליות בעבר ולכן מתקבל מחיר גבוה מהאומד לאופציה.

מציאת אומד תחתון למחיר האופציה נמצא $\tilde{V}$ כך ש־$E[\tilde{V}] < V$. נשתמש בשיטה סטטיסטית בשם $Cross\ validation$. אלגוריתם זה זהה לחישוב החסם העליון למעט שימוש בנוסחה שונה בשלב 3. נחלק את הקודקודים הבנים של $i_1 i_2, \dots, i_{n-1}$ שעבורם יש לחשב את $V_{t_n}^{i_1 i_2 \dots i_{n-1}}$ לשתי קבוצות:

1. $\frac{b}{2}$ הראשונים ישמשו לקבלת ההחלטה האם להמשיך להחזיק באופציה.
2. אם התקבלה ההחלטה להמשיך להחזיק את האופציה, $\frac{b}{2}$ האחרונים ישמשו כדי לחשב את שווי ההחזקה.

הנוסחה עבור שווי האופציה היא: $V_{t_{n-1}}^{i_1 i_2 \ldots i_{n-1}} = \left\{ \begin{array}{cc} H_{t_{n-1}}^{(B) i_1 \ldots i_{n-1}} & H_{t_{n-1}}^{(A) i_1 i_2 \ldots i_{n-1}} > (k - S_{t_{n-1}}^{i_1 i_2 \ldots i_{n-1}})_+ \\ (k - S_{t_{n-1}}^{i_1 i_2 \ldots i_{n-1}})_+ & Otherwise \end{array} \right.$.
אחרת הוא המקרה של מימוש מיידי.

נסמן את $H^{(A) i_1 i_2 \ldots i_{n-1}}$ ב : (א). הערך באיבר ה־$1,1$ במטריצה.

מתקיים: $H^{(A) i_1 i_2 \ldots i_{n-1}} = e^{-r(t_n - t_{n-1})} \frac{1}{b/2} \sum_{j=1}^{b/2} V^{i_1 i_2 \ldots i_{n-1} j}$.

עתיד מסוים עוזר לי להחליט האם להחזיק או לא.

נסמן את $H^{(B) i_1 i_2 \ldots i_{n-1}}$ ב : (ב). הערך האיבר ה־$1,2$ במטריצה. מתקיים: $H^{(B) i_1 i_2 \ldots i_{n-1}} = e^{-r(t_n - t_{n-1})} \frac{1}{b/2} \sum_{j=\frac{b}{2}+1}^{b} V^{i_1 i_2 \ldots i_{n-1} j}$.

ההבדל הוא שאת (א) נחשב בעזרת החלק הראשון, ואת (ב) לפי החלק השני.

אחרי שמחשבות את זה, מחליפות תפקידים בין שתי הקבוצות. קבוצת ההחלטה הופכת להיות קבוצת התמחור וההפך. אנחנו רוצות להגיע לשווי בנקודת הזמן.

אזי,

$V_{t_{n-1}}^{(A) i_1 i_2 \ldots i_{n-1}} = \left\{ \begin{array}{cc} H_{t_{n-1}}^{(A) i_1 \ldots i_{n-1}} & H_{t_{n-1}}^{(B) i_1 i_2 \ldots i_{n-1}} > (k - S_{t_{n-1}}^{i_1 i_2 \ldots i_{n-1}})_+ \\ (k - S_{t_{n-1}}^{i_1 i_2 \ldots i_{n-1}})_+ & Otherwise \end{array} \right.$.

לאחר חישוב שני הערכים, שווי האופציה בקודקוד יהיה הממוצע שלהם ונקבל:

$V_{t_{n-1}}^{i_1 i_2 \ldots i_{n-1}} = \frac{1}{2}(V^{(A)} + V^{(B)})$.

עם דרך חישוב חדשה זו, ממשיכות באופן איטרטיבי עד שמגיעות לראש העץ ומקבלות את האומד הנמוך $\tilde{V} = V_0$.

$E[\tilde{V}] < V$ מכיוון שאסטרטגיית המימוש שלי היא פחות מאופטימלית ונקבל את אי השוויון מאי שוויון ינסן.

חישוב הקירוב לשווי האופציה נחזור על בניית העץ וחישוב החסם העליון והחסם התחתון M פעמים ונקבל את הדגימות הבאות:

דגימות עבור החסם העליון־ $\hat{V}_1, \ldots, \hat{V}_M$.

דגימות עבור החסם התחתון־ $\tilde{V}_1, \ldots, \tilde{V}_M$.

מכאן והלאה נשתמש בשיטות הרגילות על מנת למצוא קטע בו V נמצא.

$\bar{\tilde{V}} = \frac{1}{M}\sum_{i=1}^{M}\tilde{V}_i, \bar{\hat{V}} = \frac{1}{M}\sum_{i=1}^{M}\hat{V}_i$.

לפיכך, V נמצא בקטע $\left[\bar{\tilde{V}} - \frac{\sqrt{S_n^2}}{\sqrt{M}}, \bar{\hat{V}} + \frac{\sqrt{S_n^2}}{\sqrt{M}}\right]$ - רווח סמך.

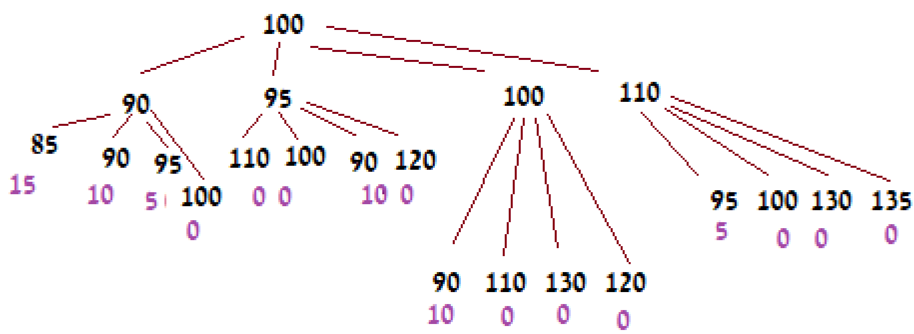
הערות 1. זמן הריצה של האלגוריתם הוא $o(Mb^n)$ כאשר n הוא מספר זמני המימוש. הוא האלגוריתם היקר ביותר שראינו בקורס ולכן הוא לא פרקטי עבור יותר מ-5 זמני מימוש.

2. לא נממש את הקוד בקורס.

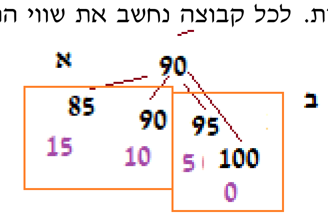
שיעור 13

דוגמא לשימוש באלגוריתם העץ המקרי

נניח שנתון העץ הבא-



ונניח שנרצה לתמחר Put ברמודית עם סטרייק $k = 100$. נניח כי הריבית היא אפס. נכתוב את שווי האופציה בסגול, כאשר שווי האופציה בתחתית העץ הוא $(k - S_T)_+$. כעת, עבור התמחור ב-t_1, יש לחלק את הענפים הבנים ל-2 . למשל עבור הקודקוד השמאלי- נחלק ל2 קבוצות. לכל קבוצה נחשב את שווי ההחזקה של האופציה (ממוצע מהוון של השווי).



$H = \frac{15+10}{2} = 12.5$

$H=\frac{5+0}{2}=2.5$

שווי האופציה בשני המקרים:

V : אם H גדול משווי המימוש המיידי, ניקח את H בתור שווי האופציה.

V: אם H גדול משווי המימוש המיידי, ניקח את H. אחרת לוקחים את שווי המימוש המיידי.

כאן, V הוא H ושווה ל2.5, מכיוון ש־$10 < 12.5$.

$V=(k-S_T)_+=10$ מכיוון ש־$10 > 2.5$.

מסקנה־ שווי האופציה בענף הנתון יהיה: $\frac{V+V}{2}=\frac{2.5+10}{2}=6.25$.

במקרה שיש 3 , עושות ממוצע על 3 החישובים. באופן דומה וניתן לחזור על התהליך עבור כל הענפים.

חוזרות על התהליך עבור 3 הענפים האחרים, חוזרות לזמן 0, וחוזרות על התהליך שוב.

אלגוריתם LSM

אלגוריתם זה שימושי לתמחור במציאות.

נשתמש בשיטת הריבועים הפחותים, Longstaff, Schwartz method רגרסיה לינארית.

רעיון האלגוריתם על מנת להעריך האם כדאי להמשיך להחזיק את האופציה מזמן המימוש t_i לזמן t_{i+1} (אופציה ברמודית)־ המטרה היא שוב לנסות להעריך את השאלה־ האם להחזיק או לממש. נניח שיש לנו את ההערכה בזמן t_{i+1} ונוכל להעריך את הערך בזמן t_i.

בהינתן שווי האופציה בזמן t_{i+1}, נשתמש ברגרסיה.

המסגרת הכללית:

בהינתן משתנה מקרי כלשהו Y, נריץ הרבה סימולציות של התהליך ונגיע למספר רב של אומדנים לשווי האופציה בזמן t_{i+1}. כעת יש לנו $V^1_{t_{i+1}}, V^2_{t_{i+1}}, ..., V^M_{t_{i+1}}$, שווי האופציה בזמן t_{i+1} ב־M סימולציות שונות.

כעת נריץ רגרסיה של Y ביחס ל־X ונוכל לבחור את X (X יכול להיות שווי הנכס בזמן t_{i+1}).

נסתכל למשל על המסלול הבא:

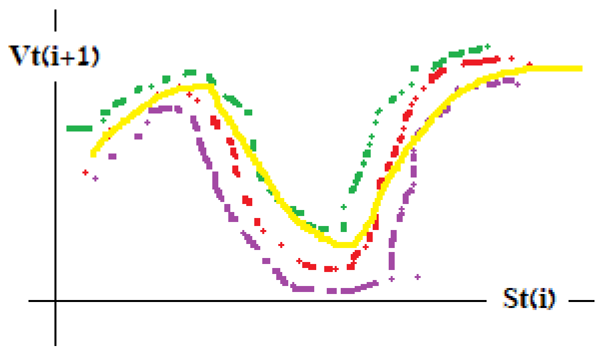
עבור $V^1_{t_{i+1}}$ יש לנו $S^1_{t_i},(S^1_{t_i})^2,(S^1_{t_i})^3$. נשים לב כי כאן הספרות העליונות הן חזקות (ולא אינדקסים).

יש לנו את ערך נכס הבסיס, ערך נכס הבסיס בריבוע וערך נכס הבסיס בשלישית בסימולציה המתאימה.

עבור $V^2_{t_{i+1}}$ יש לנו את $(S^2_{t_i}),(S^2_{t_i})^2,(S^2_{t_i})^3$.

נצייר גרף של כל הנקודות שקיבלנו־ נסתכל על שווי האופציה על ציר ה־x.

כל נקודה תהיה $(S_{t_i},V_{t_{i+1}})$ עבור כל הסימולציות שראינו.



אנחנו מחפשות פונקציה שעוברת מספיק קרוב לנקודות ושהיא יחסית פשוטה (הצהובה).

למשל, $V_{t_{i+1}}=aS_{t_i}+b(S_{t_i})^2+c(S_{t_i})^3$.

נריץ את הרגרסיה ביחס ל־3 עמודות שונות (חזקה של 1, חזקה של 2 וחזקה של 3)

למעשה, נגריל הרבה עולמות ונחפש את הקשר ביניהם ברגרסיה. כעת נרצה למצוא את המקדמים a,b,c.

בהינתן המקדמים, נקבל את הפונקציה $V_{t_{i+1}}=aS_{t_i}+b(S_{t_i})^2+c(S_{t_i})^3$ ונוכל להשתמש בה על מנת להחליט האם לממש באופן מיידי את האופציה או לא.

אלגוריתם LSM עבור תמחור $Vanilla\ put$ על נכס בסיס כלשהו S_t עם זמני מימוש אפשריים $0=t_0<t_1<...t_n=T$

שלב ראשון בוחרות פונקציות בסיס לרגרסיה (כל עמודה היא למעשה פונקציה). ברגרסיה נקבל את המקדמים $a_1,a_2,...,a_m$ ־ כאשר בחרנו m פונקציות שונות,

כך ש־$V_{t_{i+1}} = a_1 \cdot f_1(S_{t_i}) + a_2 \cdot f_2(S_{t_i}) + ... + a_m \cdot f_m(S_{t_i})$.

בדוגמא שראינו קודם לכן למשל, $f_1(x) = x$, $f_2(x) = x^2$ ו־$f_3(x) = x^3$.

במקרה בו האופציה היא על שני נכסים, נוכל לבחור את הפונקציות $S_{t_i}, S_{t_i}^2$, $X_{t_i}, X_{t_i}^2$, $S_{t_i} X_{t_i}$.

כאשר $S_{t_i}, S_{t_i}^2$ עבור הנכס הראשון, ו־$X_{t_i}, X_{t_i}^2$ עבור הנכס השני והגורם המשותף (גורם אינטראקציה) יהיה $S_{t_i} \cdot X_{t_i}$.

כאן למשל בחרנו 5 פונקציות שונות (הפונקציות פשוטות, ונרצה לקבל פונקציות פשוטות ככל האפשר).

פונקציית הרגרסיה שלנו תהיה,

$V_{t_{i+1}} = a_1 S_{t_i} + a_2 S_{t_i}^2 + a_3 X_{t_i} + a_4 X_{t_i}^2 + a_5 S_{t_i} X_{t_i}$.

כללים בבחירת פונקציות הבסיס 1. פולינומים עובדים תמיד טוב.

2. כדאי להוסיף את ערך המימוש המיידי של האופציה: $(k - S_{t_i})_+$.

3. גורמים משותפים טובים.

אין מדע מדויק מאחורי בחירת פונקציות הבסיס, וזה למעשה מהווה חיסרון של האלגוריתם.

נשים לב כי ייתכן מצב שבו "יותר מידי" פונקציות יציגו תוצאות פחות טובות.

שלב שני דוגמות M מסלולים בלתי תלויים של נכס הבסיס: $\{S_{t_0}^1, S_{t_1}^1, ..., S_{t_n}^1\}$,

$\{S_{t_0}^2, S_{t_1}^2, ..., S_{t_n}^2\}$,

.

.

$\{S_{t_0}^M, S_{t_1}^M, ..., S_{t_n}^M\}$.

החלוקה ל־n זמנים בלבד משום שנרצה לדעת את שווי נכס הבסיס רק בזמני המימוש האפשריים (נניח שהם בדידים).

לאחר יצירת המסלולים נעבוד מהסוף להתחלה.

שלב שלישי נתחיל מהסוף להתחלה־ נתמחר את האופציה מהסוף להתחלה בכל סימולציה.

בזמן הפקיעה, שווי האופציה הוא ערך המימוש המיידי.

נקבל M ערכים, $V_{t_n}^j=(k-S_{t_n}^j)_+$ עבור $j=1,2,...,M$. כאשר n הוא זמן הפקיעה.

שלב רביעי נריץ רגרסיה של אוסף הנקודות $(S_{t_{n-1}}^j, e^{-r(t_n-t_{n-1})}V_{t_n^j})$ כאשר $S_{t_{n-1}}^j$ הוא ה־X, ו־$e^{-r(t_n-t_{n-1})}V_{t_n}$ הוא ערך ה־Y - הוא למעשה הערך שחושב בשלב השלישי מהוון.

הרגרסיה תהיה על פונקציות הבסיס והנקודות יהיו אוסף הנקודות שראינו בגרף קודם לכן.

מאחר והתמחור הוא מהסוף להתחלה, המעבר הוא לא להתחלה אלא לזמן מימוש אחד אחורה.

מקבלות בכל שלב פונקציה $C_{t_{n-1}}(x)$, נכניס לה נעלמים ונקבל הערכה האם שווה להמשיך או לא.

בשלב 4 כל הסימולציות ביחד מתאחדות על מנת ליצור את הפונקציה ובשלב 5, הן שוב נפרדות.

שלב חמישי נשתמש בשלב זה בפונקציה.

עבור כל אחת מהסימולציות, נשווה בין שווי המימוש המיידי בזמן t_{n-1}: $(k-S_{t_{n-1}}^j)_+$ ובין $C_{t_{n-1}}(S_{t_{n-1}})$ (במקום ה־H באלגוריתם של העץ).

אם $C_{t_{n-1}}(S_{t_{n-1}}) < (k-S_{t_{n-1}}^j)_+$ כלומר, מימוש מיידי עדיף, אזי, $V_{t_{n-1}}^j=(k-S_{t_{n-1}}^j)_+$, שווי האופציה יהיה ערך המימוש המיידי.

אחרת, $V_{t_{n-1}}^j=e^{-r(t_n-t_{n-1})}V_{t_n}^j$. אם כדאי להחזיק את האופציה־ השווי הוא ערך האופציה בזמן המימוש הבא (מהוון). כלומר, אם יוצא שהפונקציה C "אומרת" לי להמשיך להחזיק, אני מציבה את העתיד המסלול.

סיכום

שלב שישי חוזרות על התהליך באופן איטרטיבי (עבור t_{n-2} ואז עבור t_{n-3} וכן הלאה).

נוסחת האיטרציה היא:

$$V_{t_{k-1}}^j=\left\{\begin{array}{cc}(k-S_{t_{k-1}}^j)_+ & (k-S_{t_{k-1}}^j)_+>C_{t_{k-1}}(S_{t_{k-1}}^j)\\ e^{-t(t_k-t_{k-1})}V_{t_k}^j & Otherwise\end{array}\right\}.$$

נשים לב כי פונקציית הרגרסיה משתנה־ היא ייחודית לזמן t_{k-1}. כלומר, צריך לחשב אותה מחדש ככל שעובר הזמן.

שלב שביעי לבסוף מקבלות אומד לשווי האופציה־ $\bar{V}=\frac{1}{M}\sum_{j=1}^{M}V_{t_0}^{j}$.

הערות 1. לבחירת פונקציות הבסיס של הרגרסיה יש השפעה מכרעת על דיוק התמחור.

2. מאחר וביצענו רגרסיה, ולא מצאנו את הפונקציה "האמיתית", ההחלטה האם להמשיך להחזיק באופציה אינה אופטימלית ולכן האומד המתקבל מוטה מטה, $E[\bar{V}]<V$.

3. על מנת לאזן את ההטיה מטה, נכניס ידע מהעתיד על־ידי שימוש בהמשך המסלול ה־j, $V_{t_k}^{j}$־ מופיע בנוסחה של $V_{t_{k-1}}^{j}$ והוא המאזן את ההטיה מטה.

4. אפקט ההטיה מטה (2) חזק יותר מההטיה מעלה (3) ולכן נקבל אומד מוטה מטה (רוב הזמן).

5. מציאת אומד מוטה מעלה־ מוצאות בשיטה דומה (על־ידי רגרסיה בין שלב לשלב) של הבעיה הדואלית (לא בחומר).

6. סיבוכיות זמן הריצה היא $o(b\cdot M)$ כאשר M הוא מספר הסימולציות ו־b הוא מספר פונקציות הבסיס. מה שצריך להבין מזה הוא שכל עוד מכניסים עוד פונקצית בסיס אחת־ זה לא עולה הרבה, הביטוי הוא לינארי ב־b.

7. מימוש האלגוריתם נמצא בקובץ LSM.

נבחר מימוש כל רבעון (אפשר באופן דומה להחליף את 4 ב252 כלומר, כל יום, והאלגוריתם יעבוד טוב).

מימוש Z נבחר להיות M פעמים־ M שורות, $N-1$ קפיצות, ו2 רעשי רקע שונים. כך נקבל מטריצה תלת מימדית. כלומר 2 מטריצות שונות $M\cdot(N-1)$.

בשלב הראשון נייצר מסלולים של רעשי הרקע לפי צולסקי, מהחלק השני נתחיל לתמחר את האופציה.

$S1[:,-1]$ ־ ניקח את הערך האחרון (נתחיל מהסוף להתחלה בלולאה).

על מנת לעשות רגרסיה של הנקודות (x_i,y_i) עבור $i=1,2,...,M$ על פונקציות הבסיס $f_1(x),f_2(x),...,f_b(x)$: יש לבנות מטריצה:

$$\begin{pmatrix} f_1(x_1) & f_2(x_1) & . & . & . & f_b(x_1) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ f_1(x_M) & f_2(x_M) & . & . & . & f_b(x_M) \end{pmatrix}$$

כאשר החוקיות היא שהעמודות הן הפונקציות של הבסיס והשורות הן ערכי x של הנקודות השונות. במקרה שלנו. השורות הן הסימולציות השונות.

מכניסות את ערכי y_i לוקטור $\begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_M \end{pmatrix}$ ופותרות את המערכת $\begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_M \end{pmatrix} = D \begin{pmatrix} x_1 \\ . \\ . \\ . \\ . \\ x_b \end{pmatrix}$

.

קיימת פקודה המחזירה את הפרמטרים ויש לקחת את האיבר הראשון ברשימה ([0]), משום שהיא מוסיפה גורמים לא רלוונטים עבורנו.

לאחר מכן, נכפול את הפרמטרים לפי הסדר בפונקציות.