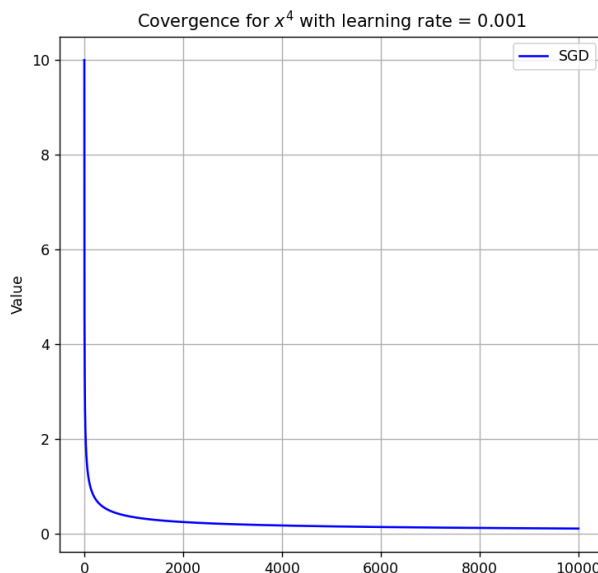
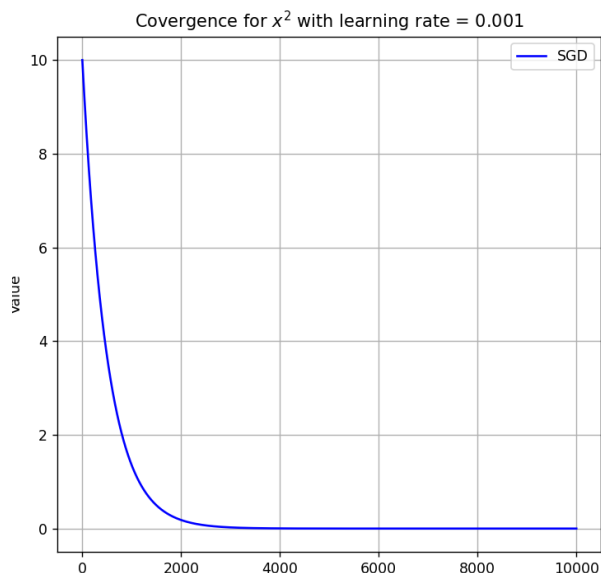


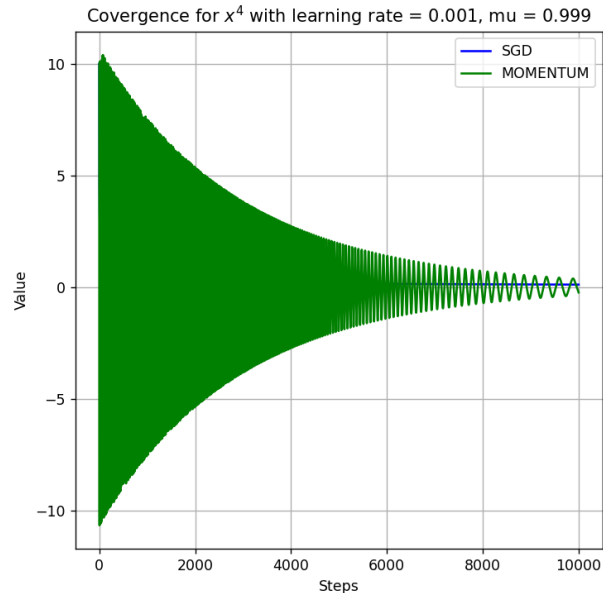
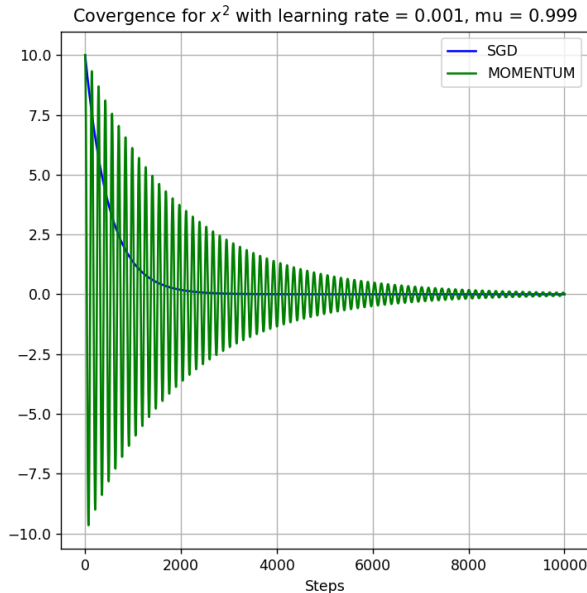
דו"ח תרגיל בית 2**תיאור העבודה:**

קיבלנו קוד למימוש של 4 פונקציות שונות למציאת מינימום של שתי פונקציות באמצעות גארדיאנט. מימשי את הפונקציות הנדרשות תוך כדי התאמות לפונקציות שהיו בטמפלייט והוספת משתנים. לאחר מכן ביצעתי בדיקות לריצה ולתוצאות ע"מ לראות את ההתכנסות של כל פונקציה למציאת המינימום תוך התאמה של ה learning rate ו momentum כדי לראות תוצאות מספיק טובות ויציבות עבור כל אחת מהשיטות.

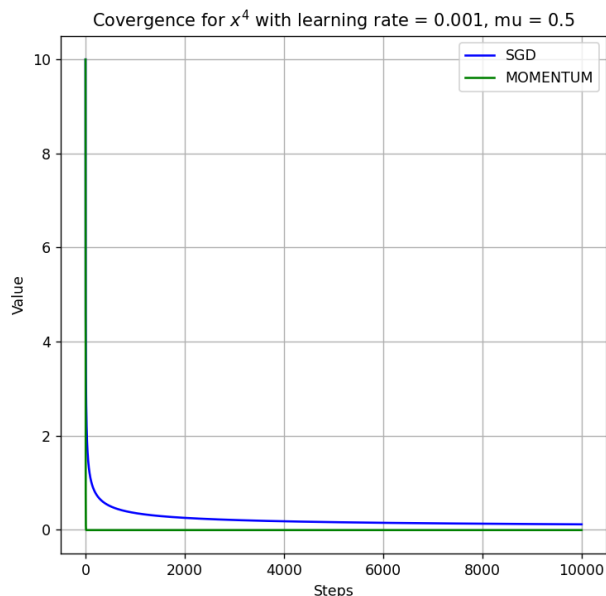
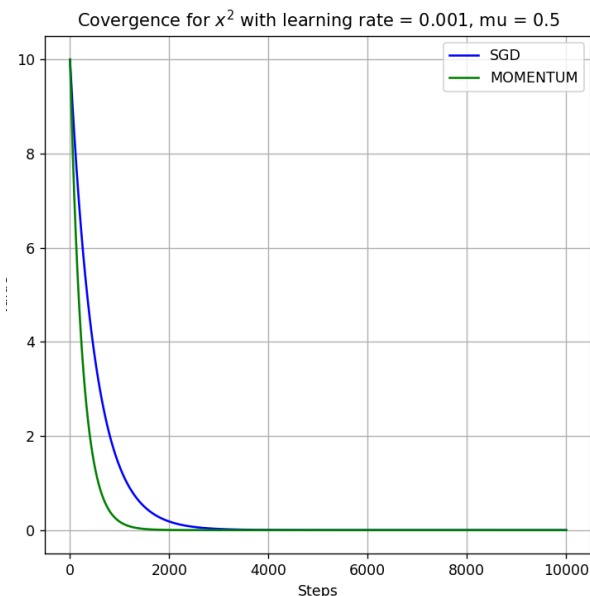
0. מימוש בוצע בקובץ הפיתון – בתחילה הרצתי עם learning rate של 0.1 וניתן היה לראות שעבור הפונקציה של x^2 הגרדיאנט הגיע במהירות ל 0, אך עבור הפונקציה x^4 , הגרדיאנט לא עבד כלל, זה הגיע למצב שהריצה קרסה מאחר ומהספרים היו גדולים או קטנים (לכיוון השלילי) מידי.
- שיחקתי קצת עם ה learning rate וראיתי שעבור learning rate של 0.001 הגרדיאנט מתכנס עבור שתי הפונקציות, בגרף למטה ניתן לראות שעבור x^4 הגרדיאנט מתקדם מהר יותר מאחר ותוצאת הנגזרת שלו גדולה יותר ולכן הקצב לימוד גדול יותר, אך ניתן לראות שעבור הפונקציה x^2 הגרדיאנט אמנם התקדם לאט אך הגיע קרוב יותר ל 0, ככה שקצב הלימוד האיטי שלו עזר לו להתקב יותר למינימום.



1. מימוש הפונקציה של המומנטום גם כן נמצאת בקוד פייתון, בתהלה שמת'י ממונטום כך שיהיה משלים ל 1 עם ה learning rate כמו שלמדנו בכיתה, כאשר בחרתי ללכת עם ה learning rate של 0.001 שראיתי שעובד עבור שתי הפונקציות. התוצאות הראו שמומנטום עבור שתי הפונקציות גרם לתנודטיות גדולה במקום לכנס את ה X למינימום, כך שנראה שנתונים שנתתי הגראדיאנט היה טוב יותר, מהיר ויציב יותר.



לכן שיחקתי קצת עם המומנטום וראיתי שכשאר אני בוחר מומנטום של 0.5 (כלומר ממוצע עם הגאדיאנט) התוצאות משתפרות עבורו (שוב עם אותו learning rate) ומראות שהוא מגיע מהר יותר וקרוב יותר למינימום.



לגבי השאלה – ניתן בהחלט לראות השפעה של המומנטום גם לחיוב וגם לשלילה עבור שתי הפונקציות. ראינו שכשיש לו ערך גבוהה מאוד הוא משפיע חזק מידי על ה X ככה שהוא מתבדר ומתקשה יותר להכנס אך כאשר מתאימים אותו ניתן לראות שהוא משפיע לחיוב על ההתקדמות וגרם לפונקציות להתכנס למינימום מהר הרבה יותר.

2. הוספתי את שתי השיטות הנוספות adagrad ו NAG, והרצתי עם אותם learning rate ומומנטום שעבדו לי טוב בשאלה הקודמת, learning rate של 0.001 ומומנטום של 0.5 ניתן לראות שה adagrad מתקדם מאוד מאוד לאט ולא מצליח להתכנס ל 0, כל השאר פחות או יותר די קרובים כאשר נראה שעבור המקרה שלנו ה momentum הוא המהיר ביותר עבור שתי הפונקציות.

