***שאלה 2:***

*יהיה:*

*כאשר ההבדל היחיד הינו פונקציית הרווח, נבדוק האם המדיניות האופטימליות שוות:*

1. *המדיניות האופטימלית* ***זהה*** *ל2 ה MDP:*

*לפי הנלמד:*

*כעת:*

*נקבל שבחירת הפעולות a עבור כל מצב (ז"א המדיניות האופטימלית), תהיה* ***זהה***  *עבור M ועבור M’,*

*זאת מכיוון ש יחזיר אותו ערך בעת הפעלה של פונקציה לינארית כפי שהתקבל.*

1. *המדיניות האופטימלית* ***זהה*** *ל2 ה MDP:*

*שוב, נחשב:*

*המעבר המסומן בשחור נובע ממשוואות בלמן- הבחירה של הפעולה a שממקסמת את הביטוי איננה מושפעת מ , לכן קיבלנו את המשוואה האחרונה \*.*

*לפי הסעיף הקודם, המדיניות האופטימליות שוות (מכיוון שקיבלנו משוואה לינארית).*

*Graphical user interface, text

Description automatically generated*

*כפי שניתן לראות, השיטה הטבלאית הניבה תוצאות טובות יותר (בדקנו והרצנו זאת מספר פעמים.*