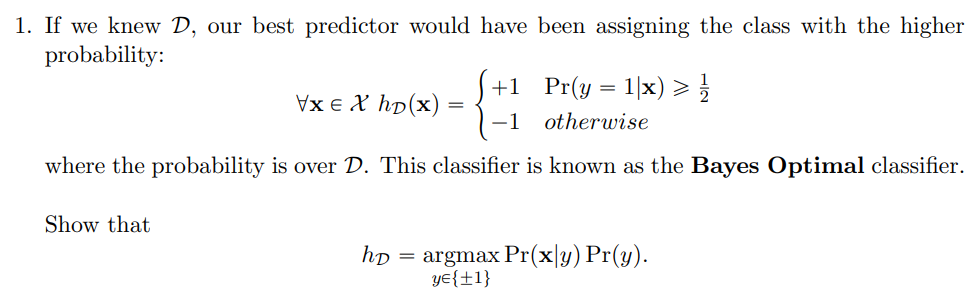
מבוא למערכות לומדות – תרגיל 3



1. ראשית ניזכר בנוסחת בייס ונוסחת ההסתברות השלמה:

נשתמש בנוסחאות ונפתח את ההגדרה עבור :

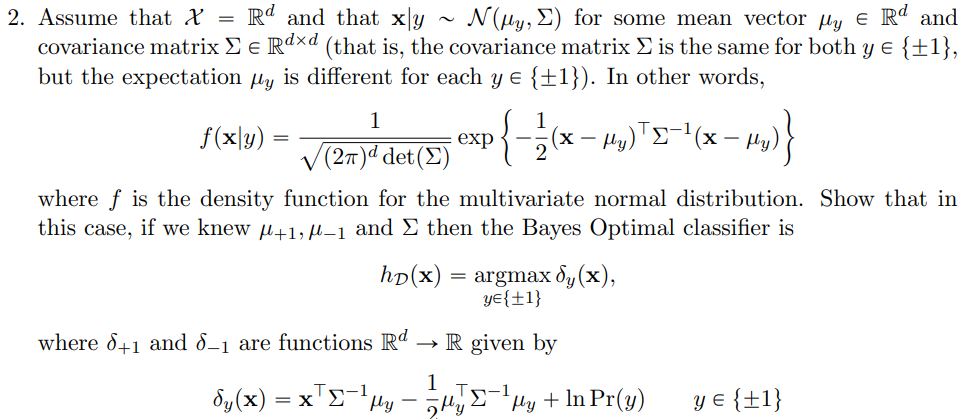
כעת נבחין בין שני מקרים אפשריים:

אם אזי

לפי הגדרת מרחב הסתברות כאשר יש רק שתי אפשרויות ל- y, נקבל חלוקה שלמה שלו ונקבל כי:

*במקרה השני*  אזי ונקבל:

*קיבלנו שההגדרה ש'ולה לכל כנדרש.*



*וכעת, מנוסחאת בייס:*

*נציב ונקבל:*

*נפעיל חוקי ln:*

*כעת שנשים לב כי הביטוי הינו קבוע ונוכל להשמיטו:*

*כעת נציב את הביטוי הנתון עבור ונקבל:*

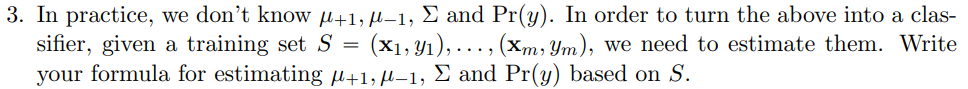
*נשים לב כי הביטוי קבוע ונוכל להשמיטו:*

*ומכיוון שהביטוי הוא קבוע:*

*נשים לב כי הביטוי הינו מטריצה סימטרית ונקבל:*

*וממונוטוניות פונקציית ה- ln נקבל:*

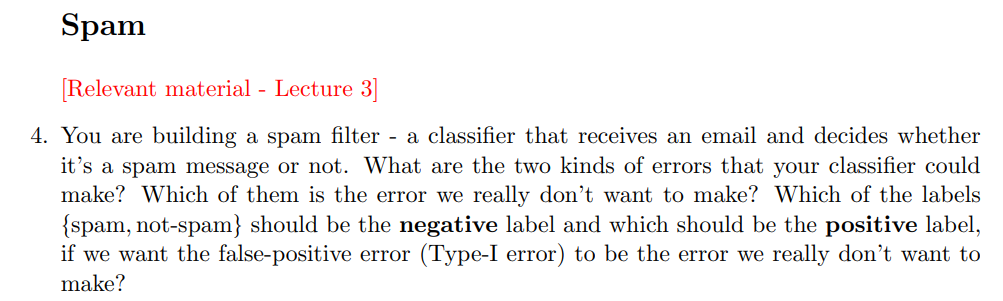
*כנדרש.*



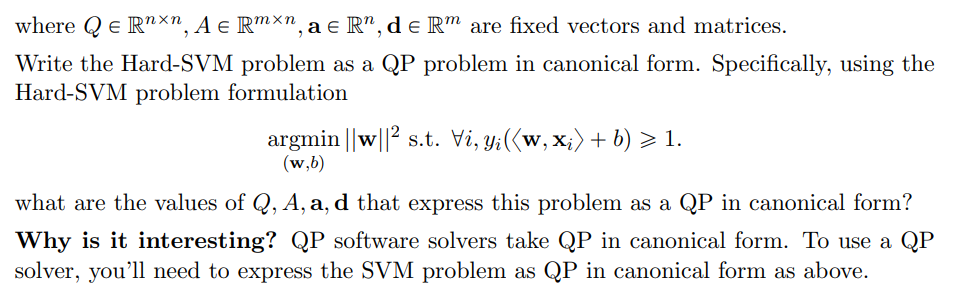
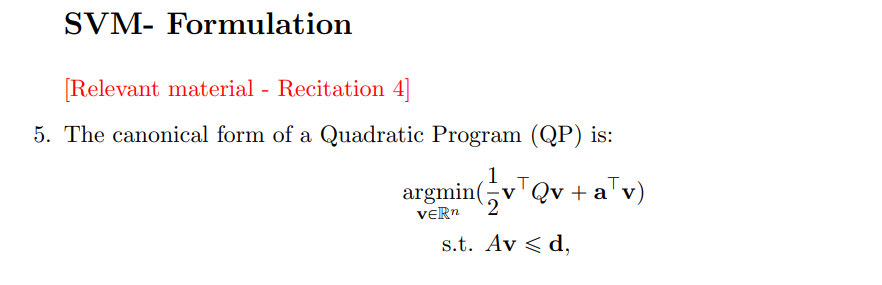
1. בקורס הסתברות למדנו כי אומד בלתי מוטה לתוחלת צריך להיות הממוצע, ובהתאם נגדיר את האומדים הנ"ל:

ולכן האומד הבלתי תלוי לשונות יהיה:

לבסוף עבור הסתברות נוכל לבחור בהסתברות ברנולי עם פרמטר p כך ש-



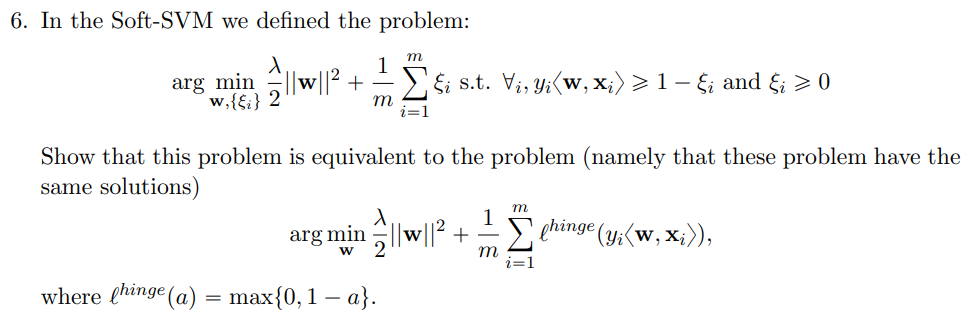
1. כשאנו רוצים לסנן מיילים לפי הפרמטר האם הם ספאם או לא נצטרך להבחין בין שתי עלויות:  
   הראשונה, כשאנו קובעים על מייל מסוים שהוא ספאם כאשר בפועל הוא אינו כזה, לכן אנו "מפספסים" משהו שהיה יכול להיות חשוב לנו.  
   השניים, שאנו מדווחים על מייל מסוים שהוא תקין אבל בפועל הוא ספאם, ובזבזנו עליו את הזמן שלנו.  
   אם כך, הטעות שנרצה להימנע ממנה יותר היא הטעות הראשונה, תיוג מייל חשוב כספאם, לכן נסמן את 1+ להיות התגית "ספאם" ו- 1- לתגית "אינו ספאם".



1. ראשית, ניזכר כי מתקיים ,

מעבר לכך, נשים לב שכל קורדינאטה מיוצגת ע"י ואנו מנסים למזער את השגיאה כמה שיותר ולכן נבחר ונקבל:

נמיר זאת לצורה הקנונית ונקבל:



1. ראשית, נשים לב כי השוני בין הביטויים נובע מהאיבר השני בחיבור ולא מהאיבר הראשון הזהה בשניהם.

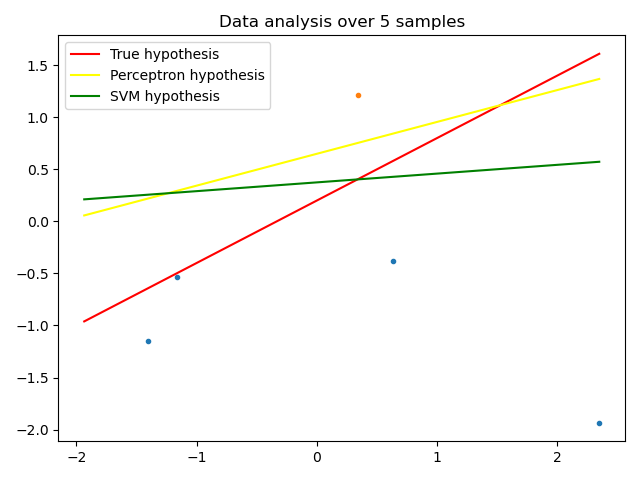
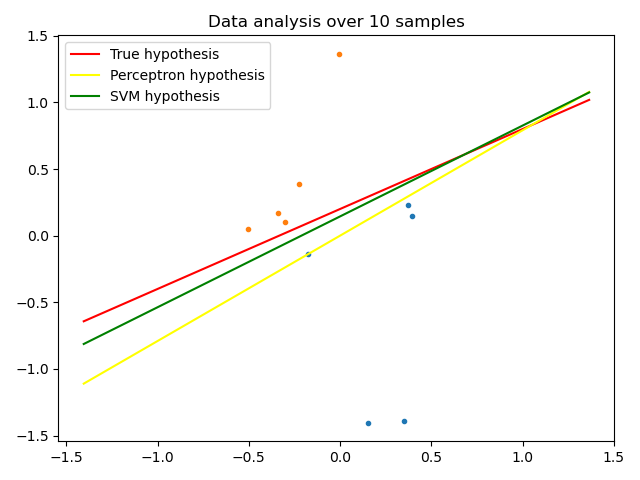
על פי הביטוי הראשון מתקיים כי

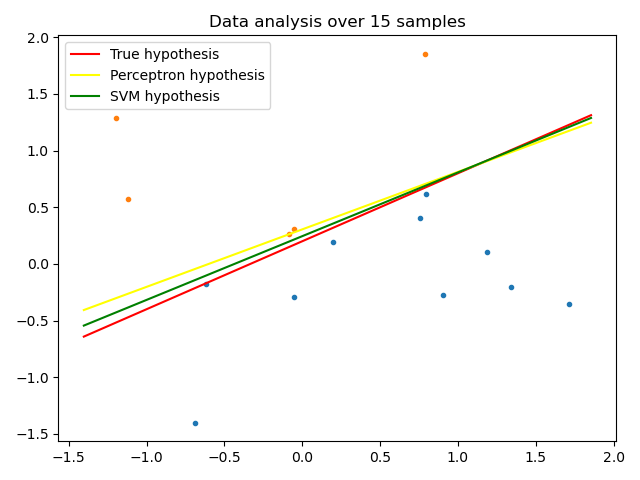
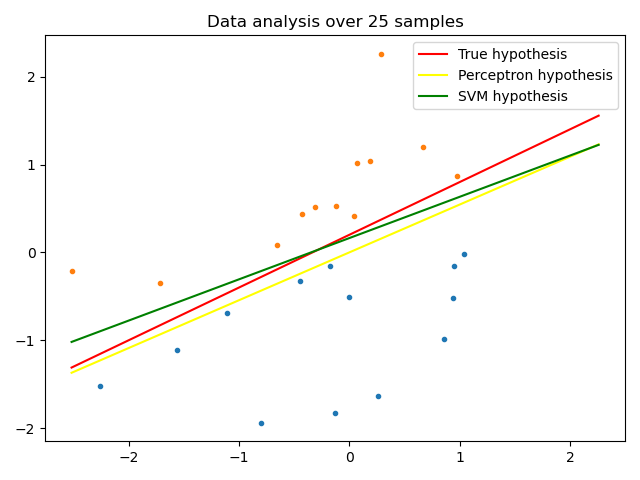
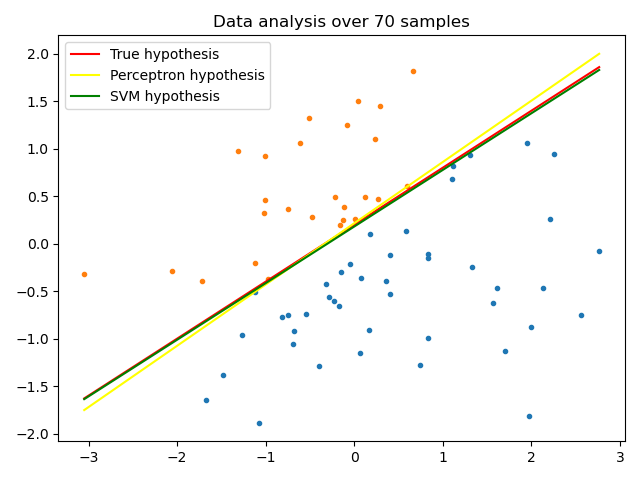
*וע"י העברת אגפים נקבל:*

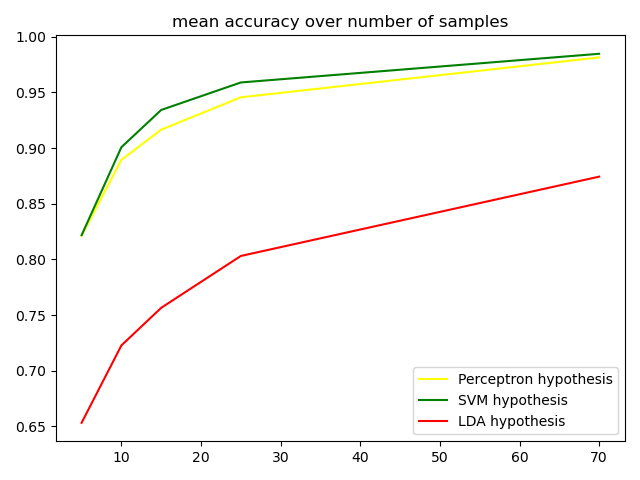
מצד שני, נתון כי לכל i, ולכן נוכל להגדיר:

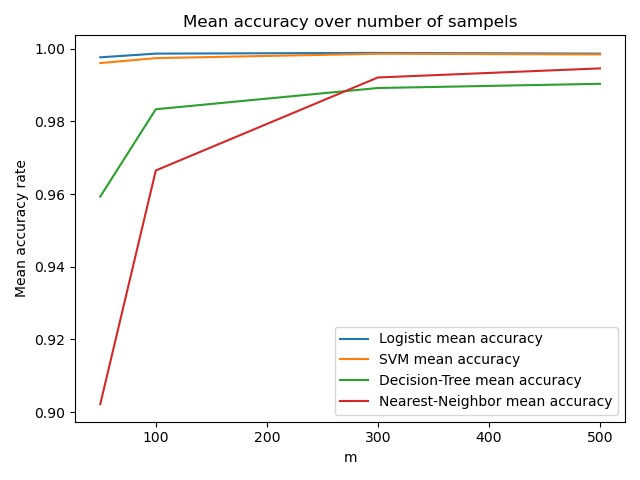
*נרצה למזער את האיבר השני בביטוי ולכן נוכל לבחור ונקבל:*

*כנדרש.*

*9.*



*10.*

12.