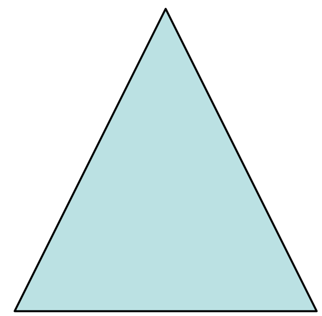
שיטת סימפלקס (Simplex):

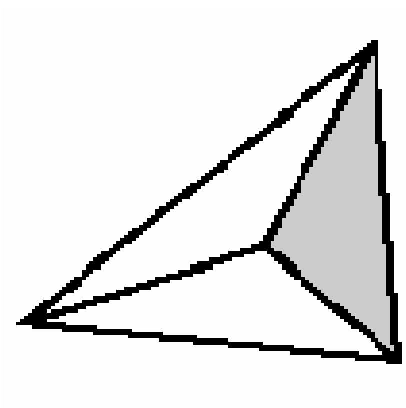
שיטה רב-מימדית עצמאית.

סימפלקס הוא הכללה של משולש. עם N משתנים, צריך N+1 נקודות בשביל להגדיר נפח ב-N מימדים בצורת סימפלקס:

N=2 :



N=3 :

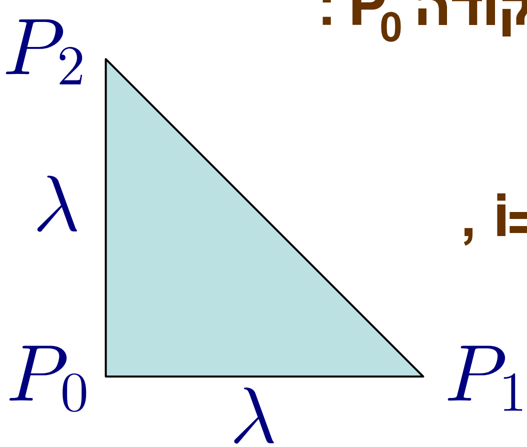


ניתן להגדיר סימפלקס התחלתי בסדר גודל λ ליד הנקודה P_0 :

$$\mathbf{P}_i = \mathbf{P}_0 + \lambda \mathbf{e}_i$$

כאשר משתמשים בווקטור יחידה e_i בכל כיוון i=1,2,…,N ,
ולוקחים את P_0 יחד עם כל ה- P_i בתור הנקודות.

למשל N=2 :

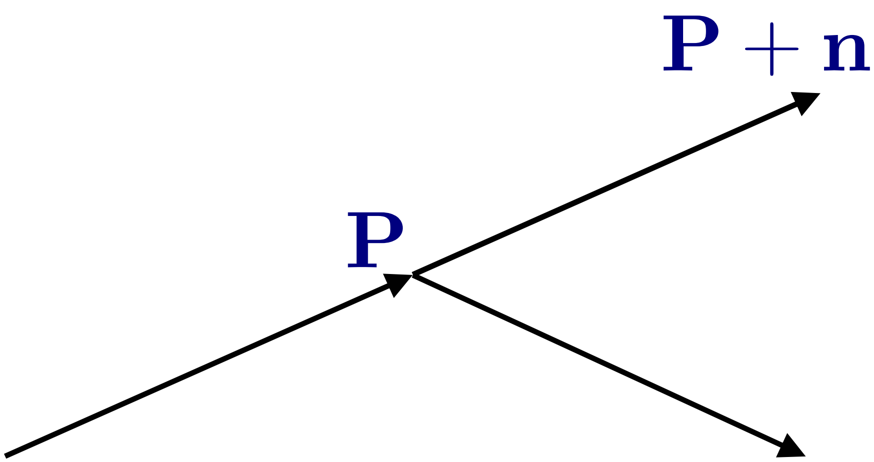


מציאת מינימום

<u>חשוב להבין:</u> N המימדים הם המשתנים שבהם תלויה הפונקציה (כאשר f עצמה לא נחשבת כמימד). הסימפלקס מטייל לו במרחב של N מימדים. לדוגמא: נגיד שרוצים למצוא את הטמפרטורה הכי נמוכה בחדר. אז יש לנו N=3 מימדים, שהם המיקום בחדר (x, y, z), ואז הפונקציה f(x, y, z) היא הטמפרטורה בנקודה (x, y, z).

הגודל λ (שקובע את מימדי הסימפלקס) נקבע על פי המרחב הנתון. למשל, בדוגמא למעלה, נבחר את λ לפי מימדי החדר (נגיד, עשירית מאורך החדר).

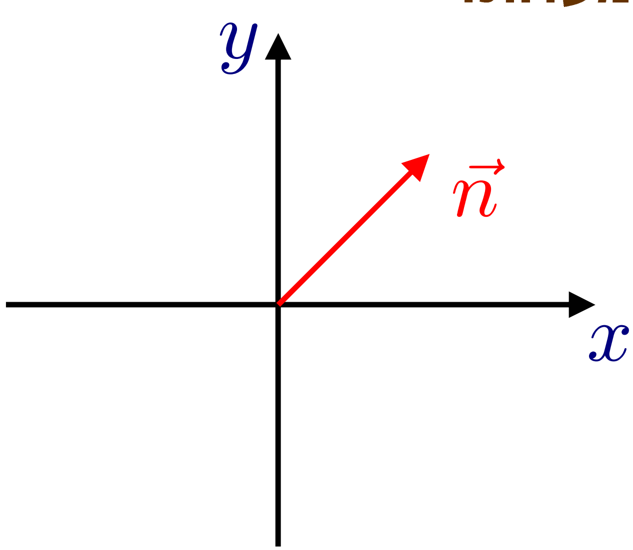
שיטות של כיוונים מצומדים Conjugate direction methods



הרעיון הכללי: סדרה של בעיות מציאת מינימום חד-מימדי, כל פעם בכיוון אחר n:

$$f(\mathbf{P} + \lambda \mathbf{n})$$

מוצאים מינימום ביחס למשתנה λ , בעזרת שיטת ברנט, למשל. אז בוחרים כיוון אחר n, וממשיכים. דוגמא עם N=2 ו- n ב- 45 מעלות:



$$f(x, y) = x^2 + y^2$$

$$\vec{n} = (1, 1)$$

בכיוון n, f היא פונקציה של משתנה אחד λ :

$\lambda \vec{n} = \lambda(1, 1)$ ז"א: $x = \lambda, \quad y = \lambda$

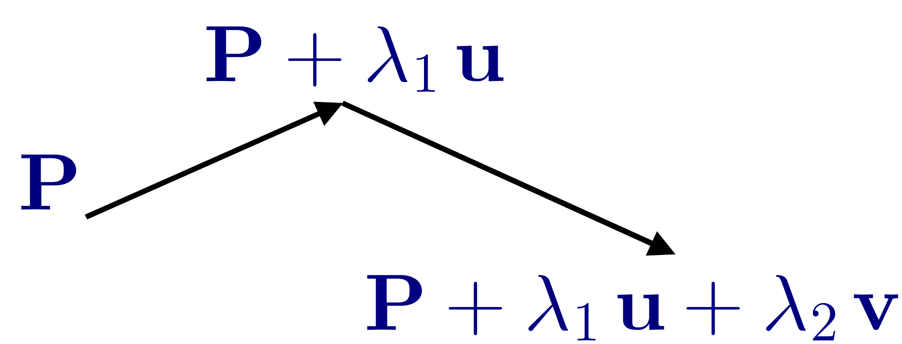
ואז הפונקציה החד-מימדית היא:

$$f(\lambda) = \lambda^2 + \lambda^2 = 2\lambda^2$$

מציאת מינימום

<u>כיצד לבחור את הכיוונים?</u> הרעיון: לא רוצים שמינימיזציה בכיוון אחד, v, תהרוס את המינימיזציה שעשינו קודם בכיוון u, כי אז נצטרך לחזור שוב ושוב על אותם כיוונים. למשל, אם הכיוונים הם:

אז יכול להיות מצב שאחרי מינימיזציה ב- u ואז ב- v:



נצטרך שוב לעשות מינימיזציה בכיוון u, אח"כ שוב ב- v, וחוזר חלילה. כך לעולם לא נסיים.

עדיף לעשות פעם אחת בכיוון u, פעם ב- v בלי להרוס את הכיוון u, ואז בכיוון w שלא הורס את u ו- v, ולהמשיך כך, ואז אחרי N כיוונים סיימנו.

מציאת מינימום

איך לקבוע את v? נתחיל עם טור טיילור סביב P ב-N משתנים, עד סדר שני:

כאן הגדרנו את האפסה של הקואורדינאטות בנקודה P=0 .

$$f(\mathbf{x}) = f(\mathbf{P}) + \sum_i \frac{\partial f}{\partial x_i}\Big|_{\mathbf{P}} x_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}\Big|_{\mathbf{P}} x_i x_j$$
$$= c - \mathbf{b}\cdot\mathbf{x} + \frac{1}{2}\mathbf{x}\cdot\mathbf{A}\cdot\mathbf{x}$$

מגדירים קבוע c, ווקטור b, מטריצה A:

ז"א:

$$\mathbf{b}_i = -\frac{\partial f}{\partial x_i}\Big|_{\mathbf{P}} \qquad \mathbf{A}_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}\Big|_{\mathbf{P}}$$

אז הגרדיינט :

$$\frac{\partial}{\partial x_k}\sum_i b_i x_i = \sum_i b_i \frac{\partial x_i}{\partial x_k} = \sum_i b_i \delta_{ik} = b_k$$

$$\delta_{ik} = \begin{cases} 1 & \text{אם } i=k: \\ 0 & \text{אם } i\neq k: \end{cases}$$

$$\frac{\partial}{\partial x_k}\sum_{i,j} A_{ij} x_i x_j = \sum_{i,j} A_{ij}(\delta_{ik} x_j + x_i \delta_{jk}) =$$

ובגלל ש-A סימטרית:

$$\sum_j A_{kj} x_j + \sum_i A_{ik} x_i = 2\sum_j A_{kj} x_j$$

מציאת מינימום

לכן:

$$\nabla f = \mathbf{A} \cdot \mathbf{x} - \mathbf{b}$$

עכשיו, אם x משתנה ב- δx, אז השינוי בגרדיינט הוא:

$$\delta(\nabla f) = \mathbf{A} \cdot \delta\mathbf{x}$$

המטרה שלנו היא למצוא מינימום, שבו: $\nabla f = \vec{0}$

נגיד שעשינו מינימיזציה בכיוון $\vec{u}$, ז"א אין ל- f נגזרת בכיוון זה:

$$\vec{u} \cdot \nabla f = 0$$

עכשיו עושים מינימיזציה בכיוון $\vec{v}$, ורוצים שהגרדיינט לא יקבל רכיב חדש בכיוון $\vec{u}$:

$$0 = \mathbf{u} \cdot \delta(\nabla f)$$

שאומר: $0 = \mathbf{u} \cdot \mathbf{A} \cdot \mathbf{v}$ בגלל ש: $\delta\mathbf{x} = \lambda\mathbf{v}$ ובאופן כללי $\lambda \neq 0$

מציאת מינימום

<u>הערה נוספת:</u> אם מגדירים את f בתור פונקציה חד-מימדית לאורך קו מסוים:

$$f(\mathbf{P}+\lambda\mathbf{n})$$

אז הנגזרת החד-מימדית: $\frac{df(\lambda)}{d\lambda}$

קשורה לגרדיינט של f הרב-מימדית בכיוון n. הנגזרת החד-מימדית היא בדיוק:

$$\vec{n}\cdot\nabla f$$

למשל, בדוגמא משקף 3: $\vec{n}=(1,1)$ $\quad$ $\nabla f=(2x,2y)$

ואז באופן כללי: $\vec{n}\cdot\nabla f=2x+2y$

אבל לאורך הקו: $y=\lambda$, $x=\lambda,$ ולכן: $\vec{n}\cdot\nabla f=4\lambda$

שזה בדיוק מה שמקבלים משקף 3 אם מחשבים את: $\frac{df(\lambda)}{d\lambda}$