10/03/2021

OPTIMIZASYON TEKNIKLERİ

Optimizasyon = En iyileme

Fayda tipli kararlarda maksimizasyon

Maliyet tipli kararlarda Minimizasyon

$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \dots \mathbb{R}$: n Boyutlu reel sayı uzayı

$X = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$: n boyutlu vektör

$f: \mathbb{R}^n \to \mathbb{R}$

$x \to f(x)$: Reel değerli fonksiyon

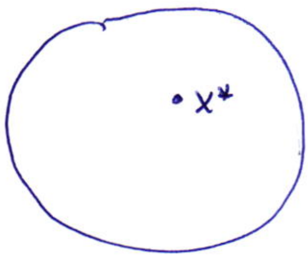
$\mathcal{X} \subset \mathbb{R}^n$ bir alt kümesi olmak üzere

Genel bir optimizasyon problemi

Amaç $\text{Max}(\text{Min})\, f(x)$

Kısıtlar $x \in \mathcal{X}$

yapısındadır.



$X^* = (x_1, x_2, \dots x_n)$

max nokta olduğunu nasıl anlarız? küme içinde sonsuz nokta var. Noktaları tek tek fonksiyonda yerine yazıp değerlerine bakıp hangisi max'ı verirse x^* odur deriz.

Tanım: $\alpha_i \in \mathbb{R}$ $(i = 1, 2, \dots, k)$ ve $x_1, x_2, \dots x_k \in \mathbb{R}^n$ olmak üzere $\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k = \sum_{i=1}^{k} \alpha_i x_i$ ifadesine $x_1, \dots, x_k$ vektörlerinin bir lineer toplamıdır. denir.

Tanım: $\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k = 0$ eşitliğini ancak ve ancak $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ için gerçekleşir ise $x_1, x_2, \dots, x_k$ vektörlerine lineer bağımsız vektörler denir. Ve $\{x_1, x_2, \dots, x_k\}$ kümesine de lineer bağımsız küme denir.

Vektörlerden biri diğeri cinsinden yazılabiliyorsa lineer bağımlı; yazılamıyorsa vektörler Lineer bağımsız olurlar.

Tanım: $x = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k$ eşitliğini sağlayan en az bir $\alpha_i \neq 0$ varsa x vektörü, $x_1, x_2, \dots, x_k$ vektörlerinin lineer toplamıdır. (kombinezonudur)

Teorem x vektörü aralarında lineer bağımsız $x_1, x_2, \dots, x_k$ vektörlerinin lineer kombinezonu olarak yazılırsa bu yazılış tek türlüdür.

İspat olmayana ergi yöntemini kullanalım. Yazılış tek türlü olmasın

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k$$

$$x = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

şeklinde en az iki farklı şekilde yazılsın. Bu iki eşitliği taraf tarafa çıkartırsak

$$0 = (\alpha_1 - \beta_1)x_1 + (\alpha_2 - \beta_2)x_2 + \dots + (\alpha_k - \beta_k)x_k$$

elde edilir. $x_1, \dots, x_k$ lineer bağımsız olduğundan

$$\alpha_1 - \beta_1 = 0, \quad \alpha_2 - \beta_2 = 0, \dots, \quad \alpha_k - \beta_k = 0$$

olmak zorundadır. Çelişki! Farklı iki yazılış olsun demiştik fakat farklı olmadı. Bu çelişkiye tek türlü yazılış olmasın kabulünden düştük. Öyleyse yazılış tek türlüdür.

Tanım: X vektör uzayının bütün elemanlarını doğuran lineer bağımsız vektörlerin oluşturduğu kümeye X'in bir tabanı denir.

$\boxed{3}$ $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ $\mathbb{R}^2$de bir tabandır. $\mathbb{R}^2$de başka bir taban yok mu?

$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ taban olsun. Eğer tabansa herhangi bir $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ vektörü (elemanı $\in X$) taban elemanları cinsinden yazılır.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta \\ \beta \end{pmatrix} \Rightarrow \begin{matrix} x_1 = \alpha_1 + \beta \\ x_2 = \beta \quad \alpha_1 = x_1 - x_2 \end{matrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 - x_2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 + x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Demekki taban tek değil.

<u>Teorem</u> Bir vektör uzayının herhangi iki tabanında eşit sayıda vektör vardır.

<u>İspat</u> Eşit sayıda olmasın ohalde $n < m$ olacak şekilde en az $\{x_1, x_2, \dots x_n\}$ ve $\{y_1, y_2, \dots y_m\}$ gibi iki farklı taban olsun. Her vektör taban vektörleri cinsinden yazılabileceğinden y_1 vektörünü diğer taban vektörleri cinsinden yazalım.

$$y_1 = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \quad \alpha_1 \neq 0 \text{ olsun.}$$

$$x_1 = \frac{1}{\alpha_1} y_1 - \frac{\alpha_2}{\alpha_1} x_2, \dots, -\frac{\alpha_n}{\alpha_1} x_n \quad \text{Bu durumda}$$

$(y_1, x_2, \dots, x_n)$ taban oluyor.

$y_2 = \beta_1 y_1 + \beta_2 x_2 + \dots + \beta_n x_n$ yazılsın.

4

y_1, y_2 aynı tabana ait elemanlardır. Bu durumda aralarında lineer bağımsızdır. Öyle ise $\beta_1 \neq 0$

$$x_2 = -\frac{\beta_1}{\beta_2} y_1 + \frac{1}{\beta_2} y_2 - \frac{\beta_3}{\beta_2} x_3 - \cdots - \frac{\beta_n}{\beta_2} x_n$$

şeklinde yazılır. Öyleyse $\{y_1, y_2, x_3, \dots, x_n\}$ de bir taban olur. İşleme bu şekilde devam edilirse $\{y_1, y_2, \dots y_n\}$ bir taban olarak bulunur. Oysa bizim $\{y_1, y_2, \dots, y_M\}$ gibi bir tabanımız vardı.

$y_{n+1} = \gamma_1 y_1 + \gamma_2 y_2 + \dots + \gamma_n y_n$ $\quad \gamma_i \neq 0$ yazıldığında

taban elemanları arasında lineer bağımsızlık kaybolacağından y_{n+1} vektörü tabana ait olmamış olur. Bu çelişkiye taban vektörlerindeki vektör sayılarının farklı olması kabulünden düştük. Öyleyse vektör sayıları eşittir.

Tanım Tabandaki lineer bağımsız vektör sayısına boyut denir.

Konveks kümeler

K bir vektör uzayının alt kümesi ve $0 \le \lambda \le 1$ reel sayı olmak üzere $\forall x, y \in K$ ve $\forall \lambda \in \mathbb{R}$ için

$\lambda x + (1-\lambda) y \in K$ ise K'ya konveks küme denir.

Konveks fonksiyonlar

Tanım: $X \subset \mathbb{R}^n$ konveks bir küme $\forall x, y \in X$ $0 \le \alpha \le 1$ olmak üzere

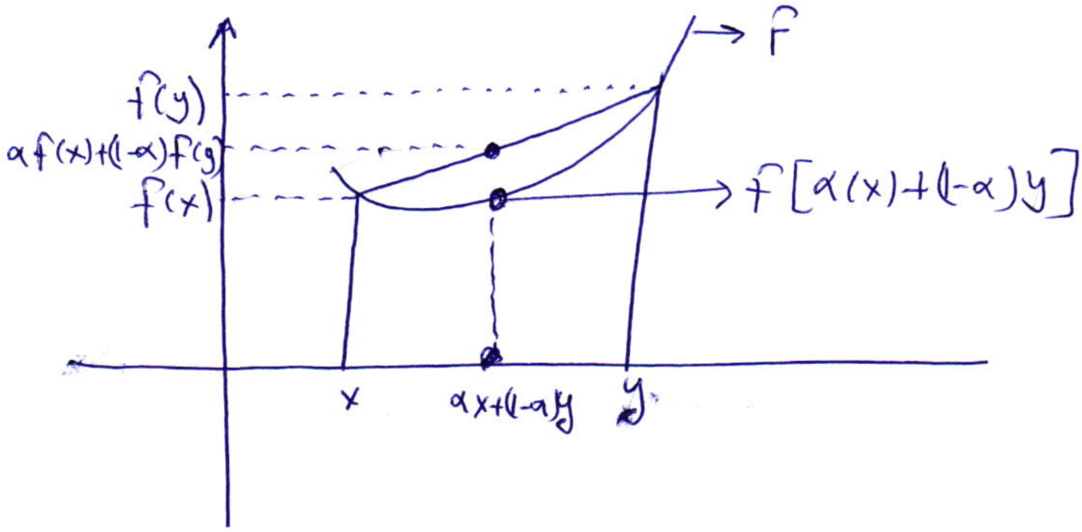
$$f[\alpha x + (1-\alpha) y] \le \alpha f(x) + (1-\alpha) f(y)$$

eşitsizliği sağlanıyorsa f'ye X'de bir konveks fonksiyon denir.

$\boxed{5}$ Tanımda $\leq$ yerine $<$ gelirse f kesin konveks

" $\geq$ gelirse f konkav

$>$ " f kesin konkav



NOT: f X'de konveks ise $-f$ konkavdır.

Teorem X bir konveks küme ve f X üzerinde en az birinci mertebeden türevleri sürekli bir fonksiyon olsun. f'nin X'de konveks olabilmesi için gerek ve yeter koşul

$$\forall x,y \in X,\ x \neq y \text{ için } \frac{f(y)-f(x)}{y-x} \geq \nabla f(x)$$

eşitsizliğinin sağlanmasıdır.

Teorem: X konveks kümesi üzerinde tanımlı f fonksiyonu en az ikinci mertebeden sürekli türevlere sahip olsun. f'nin X'de konveks olması için g.v.y.ş $\nabla^2 f(x)$ hessian matrisinin nonnegatif definit olmasıdır. kesin konveks ise $\nabla^2 f(x)$ pozitif definittir.

Bir fonksiyon sürekli ise $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ dir.

6

Tek değişkenli fonksiyonlarda fonksiyon türevli ise süreklidir. diyebiliriz ama bunu iki veya daha fazla değişkenlerde söyliyemiyoruz.

Matrislerin definitliği

$\mathbb{R}^n \times \mathbb{R}^n$ de

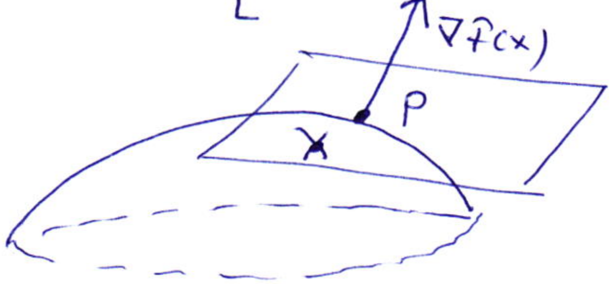
$$H = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ - & - & - & - \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{bmatrix}$$ simetrik matris olsun.

$f: \mathbb{R}^n \longrightarrow \mathbb{R}$

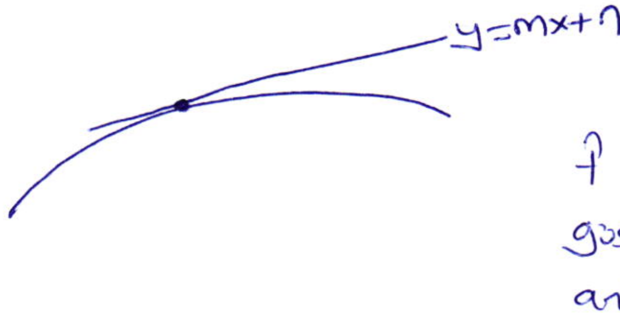
$x \longrightarrow f(x)$

$x = (x_1, x_2, \ldots x_n)$ olmak üzere

$$\nabla f(x) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots \frac{\partial f}{\partial x_n} \right]$$ vektörüne f'in gradiant vektörü denir.



yüzey üzerinde bir x noktası var. Bu noktadan yüzeye teğet düzlem çiziliyor. Bu düzlemin dik vektörü $\nabla f(x)$ dir.



lineer (doğru) ile karmaşık f fonksiyonu aynı davranışı gösterdiğinden (azalma ve artım anlamında) kolaylık açısından türevini kullanıyoruz.

ikinci Mertebeden Türevler

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

7 **1. yol** $\forall h \in \mathbb{R}^n$ ve $h \neq 0$ için

$h' H h \geq 0$ ise H non-negatif definit

$h' H h > 0$ ise H pozitif definit

$h' H h \leq 0$ ise H non-pozitif definit

$h' H h < 0$ ise H negatif definit

iii) Aksi halde H definit değildir.

2 yol $|H - \lambda I| = 0$ denkleminin çözümü ile elde

1) edilen $\lambda_1^{\geq 0}, \lambda_2^{\geq 0}, .., \lambda_n^{\geq 0}$ özdeğerleri ise non-negatif Λ (poz definit)

2) $\lambda_1 \leq 0$, $\lambda_2 \leq 0$, ... $\lambda_n \leq 0$ ise non-pozitif definit
negatif definit

3) Aksi halde definit değil.

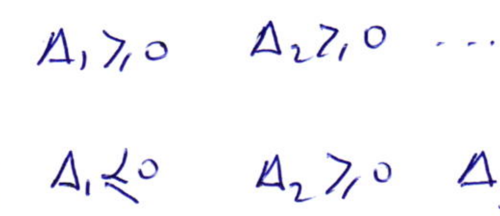
3 yol $\Delta_1 = h_{11}$ $\Delta_2 = \begin{vmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{vmatrix}$ .... $\Delta_n = \begin{vmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{vmatrix}$

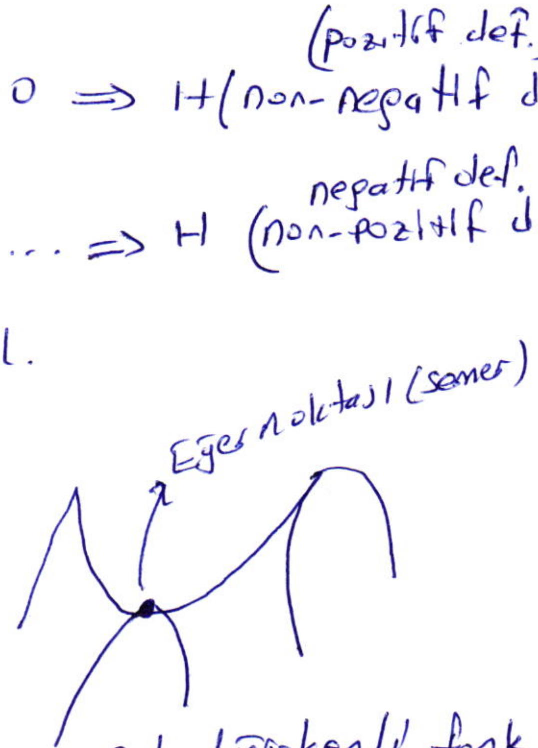
olmak üzere

1) $\Delta_1 \geq 0$ $\Delta_2 \geq 0$ ... $\Delta_n \geq 0 \Rightarrow$ H (non-negatif definit) (pozitif def.)

2) $\Delta_1 \leq 0$ $\Delta_2 \geq 0$ $\Delta_3 \leq 0$ ... $\Rightarrow$ H (non-pozitif definit) negatif def.

3) Aksi halde definit değil.





$\boxed{8}$ $H = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$ matrisinin definitliğini inceliyelim.

1 yol $\forall h \in \mathbb{R}^2$ için

$$h' H h = [h_1, h_2] \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$$

$$= [h_1, h_2] \begin{bmatrix} 2h_1 - h_2 \\ -h_1 + h_2 \end{bmatrix}$$

$$= (2h_1 - h_2)h_1 + h_2(-h_1 + h_2)$$

$$= 2h_1^2 - h_2 h_1 - h_2 h_1 + h_2^2$$

$$= h_1^2 + \underbrace{h_1^2 - 2h_1h_2 + h_2^2}_{(h_1-h_2)^2} = h_1^2 + (h_1 - h_2)^2 \geq 0$$

$\Rightarrow H$ pozitif definittir.

2. Yol $|H - \lambda I| = 0$

$$\begin{vmatrix} 2-\lambda & -1 \\ -1 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda) - 1 = 2 - 2\lambda - \lambda + \lambda^2 - 1$$

$$= 1 - 3\lambda + \lambda^2 \qquad a = 1 \quad b = -3 \quad c = 1$$

$$\Delta = b^2 - 4ac = 9 - 4 = 5$$

$$\lambda_1 = \frac{-b + \sqrt{\Delta}}{2a} = \frac{3 + \sqrt{5}}{2} \geq 0$$

$$\lambda_2 = \frac{-b - \sqrt{\Delta}}{2a} = \frac{3 - \sqrt{5}}{2} \geq 0$$

Bu durumda yine H matrisinin pozitif definit olduğu anlaşılır.

$\boxed{9}$

$y = e^x$ fonksiyonu konveks bir fonksiyon mudur?

$A = (x_1, y_1) \in X$ $\quad A = (x_1\ e^{x_1})$

$B = (x_2, y_2) \in X$ $\quad B = (x_2\ e^{x_2})$

$\alpha A + (1-\alpha)B = \alpha(x_1, e^{x_1}) + (1-\alpha)(x_2, e^{x_2})$

$$= \left(\alpha x_1 + (1-\alpha)x_2\ ,\ \alpha e^{x_1} + (1-\alpha)e^{x_2}\right)$$

$y = e^{\alpha x_1 + (1-\alpha)x_2} \neq \alpha e^{x_1} + (1-\alpha)e^{x_2}$ bundan dolayı konveks değildir.